

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2004, 11:33:18 ; Search time 80.4139 Seconds
(without alignments)
3568.829 Million cell updates/sec

Title: US-10-051-909-32
Perfect score: 4075
Sequence: 1 IRSGSLAVQTPTDLDLRR.....PLEVITEFFAVGAKQAAKA 800

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_23Sep04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	4075	100.0	800	8	ADG47936 Corn Arab
2	3774	92.6	745	8	ADM48300 Polypepti
3	3517	86.3	747	5	AAU97201 Corn Arab
4	3517	86.3	747	6	ABU08326 Corn suga
5	3517	86.3	747	8	ADG47906 Corn Arab
6	2834	69.5	740	8	ADG47941 Arabidops
7	2674	65.6	737	5	AAU97204 Soybean A
8	2674	65.6	737	6	ABU08329 Soybean s
9	2674	65.6	737	8	ADG47912 Corn Arab
10	2584	63.4	729	5	ABB93252 Herbicida
11	2584	63.4	729	8	ADG47933 Arabidops
12	2275	55.8	734	5	ABB91078 Herbicida
13	2245	55.1	729	5	ABB92622 Herbicida
14	1670.5	41.0	486	5	AAU97205 Portion o
15	1670.5	41.0	486	6	ABU08330 Soybean s
16	1670.5	41.0	486	8	ADG47914 Soybean A
17	1424	34.9	345	5	AAU97207 Portion o
18	1424	34.9	345	6	ABU08332 Wheat sug
19	1424	34.9	345	8	ADG47918 Wheat Ara
20	943	23.1	228	5	AAU97208 Portion o
21	943	23.1	228	6	ABU08333 Wheat sug
22	943	23.1	228	8	ADG47920 Wheat Ara
23	609.5	15.0	576	7	ADC07792 Rice prot
24	562	13.8	131	5	AAU97203 Portion o
25	562	13.8	131	6	ABU08328 Rice suga

26	562	13.8	131	8	ADG47910 Rice Arab
27	554.5	13.6	487	6	ABG73334 Consensus
28	554.5	13.6	487	8	ADH51665 Novel hum
29	548	13.4	488	6	ABP98504 PFAM cons
30	548	13.4	488	6	AAE35306 Human sug
31	548	13.4	488	7	ADD22918 Human sug
32	534	13.1	580	8	ADM48304 Polypepti
33	520.5	12.8	555	3	AAG29529 Arabidops
34	520.5	12.8	582	3	AAG29528 Arabidops
35	496	12.2	523	5	AAU97212 Soybean B
36	496	12.2	523	6	ABU08337 Soybean s
37	496	12.2	523	8	ADG47928 Soybean B
38	493.5	12.1	491	3	AAG32072 Arabidops
39	493.5	12.1	508	3	AAG32071 Arabidops
40	493	12.1	493	8	ADN72405 Thale cre
41	492	12.1	446	6	ABU43281 Protein e
42	489.5	12.0	513	5	AAU97210 Corn Beta
43	489.5	12.0	513	6	ABU08335 Corn suga
44	489.5	12.0	513	8	ADG47924 Corn Beta
45	489	12.0	529	5	AAU97214 Wheat sug

ALIGNMENTS

RESULT 1
ADG47936
ID ADG47936 standard; protein; 800 AA.
XX
AC ADG47936;
XX
DT 11-MAR-2004 (first entry)
XX
DE Corn Arabidopsis-like sugar transport protein #2.
XX
KW Arabidopsis-like sugar transport protein;
KW Beta-vulgaris-like sugar transport protein; transgenic; physical mapping;
KW corn; plant.
XX
OS Zea mays.
XX
PN US2002199217-A1.
XX
PD 26-DEC-2002.
XX

17-JAN-2002; 2002US-00051909.
XX
24-APR-1998; 98US-0083044P.
14-APR-1999; 99US-00291922.
XX
(HELE/) HELENTJARIS T G.
XX
Helentjaris TG;
XX
WPI; 2004-040967/04.
DR N-PSDB; ADG47935.
XX
New isolated polynucleotide encoding a polypeptide having sugar transport protein activity, for producing a transformed plant and for use as probes in physical mapping.
XX
Claim 18; SEQ ID NO 32; 71pp; English.
XX
The invention relates to Arabidopsis-like or Beta-vulgaris-like sugar transport proteins and their corresponding nucleic acid sequences. The sequences of the invention are useful to transform a cell. These are also useful to produce a transgenic plant. Probes derived from sequences encoding sugar transport protein may be used for physical mapping. The present sequence is corn Arabidopsis-like sugar transport protein.
XX
SQ Sequence 800 AA;

Query Match 100.0%; Score 4075; DB 8; Length 800;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 800; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IRSGSWLAVQTPFTPDLDLRRERLLPSVVLALPGELPPASCSSQEPVTSDDILEDKMSGAV 60
Db 1 IRSGSWLAVQTPFTPDLDLRRERLLPSVVLALPGELPPASCSSQEPVTSDDILEDKMSGAV 60

Qy 61 LVAIVASIGNLLQGWDNATIAAAVLYIKKEFQIQNEPTVEGLIVMSLIGATIVTTFSGP 120
Db 61 LVAIVASIGNLLQGWDNATIAAAVLYIKKEFQIQNEPTVEGLIVMSLIGATIVTTFSGP 120

Qy 121 LSDSIGRRPMLILSSILYFFSGLIMLWSPNVYVLLARFVDGFGIGLAVTLVPLYISEIA 180
Db 121 LSDSIGRRPMLILSSILYFFSGLIMLWSPNVYVLLARFVDGFGIGLAVTLVPLYISEIA 180

Qy 181 PSEIRGLLNTLPQFSGSGMFLSYCMVFGMSLSPSPDWRIMLGVLAIPLSFFFGLTIFYL 240
Db 181 PSEIRGLLNTLPQFSGSGMFLSYCMVFGMSLSPSPDWRIMLGVLAIPLSFFFGLTIFYL 240

Qy 241 PESPRWLVSIGRMAEAKKVLQKLRGKDDVSGELSLLEGLVGGDTSTEEYIIGPATEAA 300
Db 241 PESPRWLVSIGRMAEAKKVLQKLRGKDDVSGELSLLEGLVGGDTSTEEYIIGPATEAA 300

Qy 301 DDLVTDGDKQITLYGPEEGQSWIARPSKGPIMLGSVLSASRHGSMVNSQSVPLMDPIVT 360
Db 301 DDLVTDGDKQITLYGPEEGQSWIARPSKGPIMLGSVLSASRHGSMVNSQSVPLMDPIVT 360

Qy 361 LFGSVHENMPQAGGMRSTLFPNFGSMFSTDDQAKNEOWDEENLHRDDEYASDGAGGD 420
Db 361 LFGSVHENMPQAGGMRSTLFPNFGSMFSTDDQAKNEOWDEENLHRDDEYASDGAGGD 420

Qy 421 YEDNLHSPLLSRQATGAEGKDIVHHGHRGSAISMRQTLLEGCGDVSSTDGGGQWLAW 480
Db 421 YEDNLHSPLLSRQATGAEGKDIVHHGHRGSAISMRQTLLEGCGDVSSTDGGGQWLAW 480

Qy 481 KWSEKEGEGNKGKGFVYLVHQQEGVPGSRGSIIVSLPGGGDVFESEFVHAAALVSQSA 540
Db 481 KWSEKEGEGNKGKGFVYLVHQQEGVPGSRGSIIVSLPGGGDVFESEFVHAAALVSQSA 540

Qy 541 LFSKGLAEPRMSDAAWHPSEVAAKGRWKDLFEPGVRALLVGVGIQLQFAGINGVL 600
Db 541 LFSKGLAEPRMSDAAWHPSEVAAKGRWKDLFEPGVRALLVGVGIQLQFAGINGVL 600

Qy 601 YTPQILEQAGVAVILSKFGLSSASASILISLTLTLLMLPCIGFAMLLMDLSGRRFLLG 660
Db 601 YTPQILEQAGVAVILSKFGLSSASASILISLTLTLLMLPCIGFAMLLMDLSGRRFLLG 660

Qy 661 TIPILIASLVLVNLDLGTALHALLSTVSVIVYFCCFVMGFGPIPNILCAEIFFTRV 720
Db 661 TIPILIASLVLVNLDLGTALHALLSTVSVIVYFCCFVMGFGPIPNILCAEIFFTRV 720

Qy 721 RGLCIAICAFTHWIGDIIVTYSLPVMLNAGLAGVFSIYAVVCLISFVFLKVPETKGM 780
Db 721 RGLCIAICAFTHWIGDIIVTYSLPVMLNAGLAGVFSIYAVVCLISFVFLKVPETKGM 780

Qy 781 PLEVITEFFAVGAKQAAKA 800
Db 781 PLEVITEFFAVGAKQAAKA 800

RESULT 2

ID ADM48300 standard; protein; 745 AA.

XX

AC ADM48300;

XX

DT 03-JUN-2004 (first entry)

XX

DE Polypeptide sequence #350 useful in producing transgenic plants.

XX

KW Plant; maize; rice; sorghum; thale cress; soybean; wheat; transgenic;

KW osmotic stress; sugar transport; cell cycle pathway; plant height;

KW carbohydrate transport; crop productivity; plant growth;

KW stress resistance; disease resistance; insect resistance; heat tolerance;
KW nitrogen assimilation; water stress tolerance;
KW photosynthetic carbon fixation; virus resistance; gene therapy.

OS Zea mays.

XX US2003233670-A1.

XX 18-DEC-2003.

XX 04-DEC-2002; 2002US-00310154.

XX 04-DEC-2001; 2001US-0337358P.

XX (EDGE/) EDGERTON M D.

XX (CHOM/) CHOMET P S.

XX (LACC/) LACCETTI L B.

PI Edgerton MD, Chomet PS, Laccetti LB;

XX WPI; 2004-061374/06.

DR N-PSDB; ADM47932.

XX

PT New polynucleotide, useful for manipulating plant protein quality,
PT improving plant growth, yield and crop productivity or grain composition
PT or producing plants with improved properties.

XX

PS Claim 8; SEQ ID NO 718; 144pp; English.

XX

CC The present invention relates to polynucleotide sequences, and the
CC proteins they encode. The sequences are isolated from a variety of
CC organisms such as plants (e.g. maize, rice, sorghum, thale cress,
CC soybean, and wheat), cyanobacteria, bacteria, yeast and other fungi. The
CC polynucleotide and polypeptide sequences of the invention are useful in
CC the production of transgenic plants that have improved properties. Also
CC disclosed are methods of producing fertile transgenic plants, preferably
CC maize, with desired phenotypes. The polynucleotide and polypeptide
CC sequences are useful for improving plants by providing protection against
CC osmotic stress, improving altering sugar transport and/or metabolism,
CC modifying the cell cycle pathway, reducing plant height, modifying
CC carbohydrate transport, improving crop productivity, improving plant
CC growth and stress resistance, improving disease resistance, improving
CC insect resistance, improving cold or heat tolerance, improving nitrogen
CC assimilation, improving stalk strength, improving water stress tolerance,
CC improving photosynthetic carbon fixation, improving biotic and abiotic
CC stress resistance, improving resistance to oxidative stress, providing
CC increased vigour, reducing senescence, and conferring virus resistance.

CC The present sequence represents a polypeptide sequence of the invention.
CC Note: The sequence data for this patent is not provided in the printed
CC specification but is obtained in electronic format from the USPTO website
CC at seqdata.uspto.gov.

XX

SQ Sequence 745 AA;

Query Match 92.6%; Score 3774; DB 8; Length 745;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 742; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 56 MSGAVLVAIVASIGNLLQGWDNATIAAAVLYIKKEFQIQNEPTVEGLIVMSLIGATIVT 115
Db 1 MSGAVLVAIVASIGNLLQGWDNATIAAAVLYIKKEFQIQNEPTVEGLIVMSLIGATIVT 60

QY 116 TFSGPLSDSICRRPMLILSSILYFFSGLIMLWSPNVYVLLARFVDGFGIGLAVTLVPLY 175
Db 61 TFSGPLSDSICRRPMLILSSILYFFSGLIMLWSPNVYVLLARFVDGFGIGLAVTLVPLY 120

QY 176 ISEIAPSEIRGLLNTLPQFSGSGMFLSYCMVFGMSLSPSPDWRIMLGVLAIPLSFFFG 235
Db 121 ISEIAPSEIRGLLNTLPQFSGSGMFLSYCMVFGMSLSPSPDWRIMLGVLAIPLSFFFG 180

QY 236 TIFYLPESPRWLVSIGRMAEAKKVLQKLRGKDDVSGELSLLEGLVGGDTSTEEYIIGP 295
Db 181 TIFYLPESPRWLVSIGRMAEAKKVLQKLRGKDDVSGELSLLEGLVGGDTSTEEYIIGP 240

QY	296	ATEAADDLVTDGDK	EQITLYGPEEGQSWIARPSKGPIMLGSVLSASRHGSMVNQSVPLM	355
Db	241	ATEAADDLVTDGDK	EQITLYGPEEGQSWIARPSKGPIMLGSVLSASRHGSMVNQSVPLM	300
QY	356	DPIVTLFGSVHENMP	QAGGSMRSTLFPNFGSMFVTDQHAQNEQWDEENLHRDDEEYASD	415
Db	301	DPIVTLFGSVHENMP	QAGGSMRSTLFPNFGSMFVTDQHAQNEQWDEENLHRDDEEYASD	360
QY	416	GAGGDYEDNLHSP	LLSRQATGAEGKDIVHHGHRGSALSMMRRTLLGEGDGVSSSTDIGGG	475
Db	361	GAGGDYEDNLHSP	LLSRQATGAEGKDIVHHGHRGSALSMMRRTLLGEGDGVSSSTDIGGG	420
QY	476	WQLAWKWKSEKEG	ENGRKEGFKRVYLHQEGVPGSRGSIIVSLPGGDFVFESEFVHAAAL	535
Db	421	WQLAWKWKSEKEG	ENGRKEGFKRVYLHQEGVPGSRGSIIVSLPGGDFVFESEFVHAAAL	480
QY	536	VSQSALFSKGLAE	PRMSDAAMVHPSEVAAGSRWKDLFEPGVRRALLVGVGIQLQQFAG	595
Db	481	VSQSALFSKGLAE	PRMSDAAMVHPSEVAAGSRWKDLFEPGVRRALLVGVGIQLQQFAG	540
QY	596	INGVLYTTPQILE	QAGVAVILSKFGLSSASASILISLTTLLMPCIGFAMLLMDLSGRR	655
Db	541	INGVLYTTPQILE	QAGVAVILSKFGLSSASASILISLTTLLMPCIGFAMLLMDLSGRR	600
QY	656	FLLLGTIPILIAS	LVVSNLIDLGTLAALLSTVSVIVVFCCFVMGFGPIPNILCAEI	715
Db	601	FLLLGTIPILIAS	LVVSNLIDLGTLAALLSTVSVIVVFCCFVMGFGPIPNILCAEI	660
QY	716	FPTRVRGLCIAI	CAFTFWIGDIIIVTYSLPVMLNAIGLAGVFSIYAVVCLISFVFELKVP	775
Db	661	FPTRVRGLCIAI	CAFTFWIGDIIIVTYSLPVMLNAIGLAGVFSIYAVVCLISFVFELKVP	720
QY	776	ETKGMPLLEVIT	EFFAVGAKQAAKA	800
Db	721	ETKGMPLLEVIT	EFFAVGAKQAAKA	745
RESULT 3				
ID	AAU97201	standard; protein; 747 AA.		
XX	AC	AAU97201;		
XX	XX	27-AUG-2002	(first entry)	
DE	XX	Corn Arabidopsis thaliana-like sugar transport protein.		
XX	XX	Corn; Arabidopsis thaliana-like sugar transport protein;		
KW	XX	carbohydrate transport; grain filling; annual field crop; plant.		
OS	XX	Zea mays.		
FT	Key	Location/Qualifiers		
FT	Misc-difference	129	/label= Unknown	
FT	Misc-difference	133.134	/label= Unknown	
FT	Misc-difference	144	/label= Unknown	
FT	Misc-difference	178	/label= Unknown	
FT	Misc-difference	207	/label= Unknown	
FT	Misc-difference	218	/label= Unknown	
FT	Misc-difference	220	/label= Unknown	
FT	Misc-difference	236	/label= Unknown	
PN	XX	US6383776-B1.		

PD	07-MAY-2002.			
XX	14-APR-1999;	99US-00291922.		
XX	24-APR-1998;	98US-0083044P.		
XX	(DUPO) DU PONT DE NEMOURS & CO E I.			
PI	Allen SM, Hitz WD, Kinney AJ, Tingey SV;			
XX	WPI; 2002-453364/48.			
DR	N-PSDB; ABK51962.			
XX	New nucleic acid encoding plant sugar-transport proteins, useful for preparing transgenic plants with altered carbohydrate distribution.			
PS	Claim 1; Fig 1; 54pp; English.			
XX	The present invention relates to the isolation of plant polynucleotide sequences encoding an Arabidopsis thaliana-like sugar transport protein or Beta vulgaris-like sugar transport protein. The polynucleotide sequences are useful for altering the level of sugar transport proteins in plants, i.e. for control of carbohydrate transport and distribution in plant cells, e.g. during grain filling of annual field crops (e.g. corn, rice, soybeans, and wheat), and, for studying carbohydrate flows and sugar transport. The polynucleotide sequences can also be used to isolate cDNA sequences and genes that encode homologues of the new proteins. The present sequence represents a corn Arabidopsis thaliana-like sugar transport protein			
XX	Sequence 747 AA;			
SQ	Query Match	86.3%;	Score 3517;	DB 5; Length 747;
	Best Local Similarity	91.6%;	Pred. No. 0;	
	Matches 684;	Conservative 27;	Mismatches 34;	Indels 2; Gaps 2;
QY	56	MSGAVLVAVASIGNLLQGW	DNATIAAAVLYIKKEFQLQNEPTVEGLIVMSMLIGATIVT	115
Db	1	MGGAVMVAATAASIGNLLQGW	DNATIAAGAVLYIKKEFNQLQSEPLIEGLIVAMFLIGATVIT	60
QY	116	TFSGPLSDSIGRRPMLILSS	ILYFFESGLIMLWSPNVYVLLARFVDGFGIGLAVTLVPLY	175
Db	61	TSPGPRADCVGRRPMLVAS	AVLYFVSGLVMLWAPIVYILLARLIDGFGIGLAVTLVPLY	120
QY	176	ISEIAPSEIR-GLLNTLPQ	FSG-SGGMFLSYCMVFGMSLSPSPDWIRIMLGVALPSLFFF	233
Db	121	ISETAPHRXSWGXXNTLP	QFIGVXGCMFLSYCMVFGMSLMPKPDWRLMLGVLSPSLXYF	180
QY	234	GLTIFYLPESPRWLVSKGR	MAEAKVQLKRGKDDVSGELSLLEGLVGGDTSEIEYII	293
Db	181	GLTVFYLPESPRWLVSKGR	MAEAKRVXQRLRGREDVSEXALLVEGLGVGKDTRIXEYII	240
QY	294	GPATEAADDLVTDGDKEQ	ITLYGPEEGQSWIARPSKGPIMLGSVLSASRHGSMVNQSV	353
Db	241	GPATEAADDLVTDGDKEQ	ITLYGPEEGQSWIARPSKGPIMLGSVLSASRHGSMVNQSV	300
QY	354	LMDPIVTLFGSVHENMP	QAGGSMRSTLFPNFGSMFVTDQHAQNEQWDEENLHRDDEEYA	413
Db	301	LMDPIVTLFGSVHENMP	QAGGSMRSTLFPNFGSMFVTDQHAQNEQWDEENLHRDDEEYA	360
QY	414	SDGAGGDYEDNLHSP	LLSRQATGAEGKDIVHHGHRGSALSMMRRTLLGEGDGVSSSTDIG	473
Db	361	SDGAGGDYEDNLHSP	LLSRQATGAEGKDIVHHGHRGSALSMMRRTLLGEGDGVSSSTDIG	420
QY	474	GGWQLAWKWKSEKEG	ENGRKEGFKRVYLHQEGVPGSRGSIIVSLPGGDFVFESEFVHAA	533
Db	421	GGWQLAWKWKSEKEG	ENGRKEGFKRVYLHQEGVPGSRGSIIVSLPGGDFVFESEFVHAA	480
QY	534	ALVQSALFSKGLAE	PRMSDAAMVHPSEVAAGSRWKOLFEPGVRRALLVGVGIQLQQF	593
Db	481	ALVQSALFSKGLAE	PRMSDAAMVHPSEVAAGSRWKOLFEPGVRRALLVGVGIQLQQF	540
QY	594	AGINGVLYTTPQILE	QAGVAVILSKFGLSSASASILISLTTLLMPCIGFAMLLMDLSG	653

Db 541 AGINGVLYYTPQILEQAGVAVILSKFGLSSASASILISSLTTLMLPCIGFAMLLMDLSG 600
QY 654 RRFLLLGTIPILIASLVILVSNLIDGLTLAHALLSTVSIVVYFCFWMGFGPIPNILCA 713
Db 601 RRFLLLGTIPILIASLVILVSNLIDGLTLAHALLSTVSIVVYFCFWMGFGPIPNILCA 660
QY 714 EIFPTRVRGLCIAICAFTEFWIGDIIIVTYSLPVMLNAIGLAGVFSIYAVVCLISFVFVLK 773
Db 661 EIFPTRVRGLCIAICAFTEFWIGDIIIVTYSLPVMLNAIGLAGVFSIYAVVCLISFVFVLK 720
QY 774 VPETKGMPLVITEFFAVGAKQAAKA 800
Db 721 VPETKGMPLVITEFFAVGAKQAAKA 747

RESULT 4
ABU08326
ID ABU08326 standard; protein; 747 AA.
XX AC ABU08326;
XX

DT 29-MAY-2003 (first entry)
XX DE Corn sugar transport protein #1.
XX

KW Arabidopsis thaliana-like sugar transport protein; corn; rice; wheat;
KW plant sugar transport protein; carbohydrate transport; soybean;
KW carbohydrate distribution; plant.
XX

OS Zea mays.
XX PN US2002178468-A1.
XX

PD 28-NOV-2002.
XX

PF 17-JAN-2002; 2002US-00051902.
XX

PR 24-APR-1998; 98US-0083044P.
PR 14-APR-1999; 99US-00291922.
XX

PA (ALLE/) ALLEN S M.
PA (HITZ/) HITZ W D.
PA (KINN/) KINNEY A J.
PA (TING/) TINGEY S V.
XX

PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;
XX WPI; 2003-340957/32.
DR N-PSDB; ABX93198.
XX

PT Novel plant sugar transport proteins and nucleic acid encoding the
PT protein useful for producing transgenic plants having altered levels of
PT sugar transport protein.
XX

PS Claim 5; Fig 1; 56pp; English.
XX

CC The present invention relates to the isolation of Arabidopsis thaliana-
CC like or Beta vulgaris-like sugar transport proteins, and the
CC polynucleotide sequences encoding them. The plant sugar transport
CC proteins of the invention have been isolated from corn, rice, soybean,
CC and wheat. The polypeptides of the invention may be used for altering the
CC level of expression of a sugar transport protein in a host cell, by
CC transforming a host cell with a chimeric construct encoding all, or a
CC portion of the sugar transport protein, in sense or antisense
CC orientation. Particularly, the polypeptides may provide a means to
CC control carbohydrate transport and distribution in plants. ABU08326-
CC ABU08333 represent Arabidopsis thaliana-like sugar transport proteins
XX

SQ Sequence 747 AA;

Query Match 86.3%; Score 3517; DB 6; Length 747;
Best Local Similarity 91.6%; Pred. No. 0;

Matches 684; Conservative 27; Mismatches 34; Indels 2; Gaps 2;
QY 56 MSGAVLVAIVASIGNLLQGDNDATIAAVALYIKKEFQLONEPTVEGLIVSMSLIGATIVT 115
Db 1 MCGAVMVAIAASIGNLLQGDNDATIAAGVALYIKKEFNLOSEPLIEGLIVAMFLIGATVIT 60
QY 116 TFGPLSDSIGRRPMLILSSILYFFSGLIMLWSPNVVLLARFVDGFGIGLAVTLVPLY 175
Db 61 TSPGPRADCVGRRPMLVASAVLYFVSGLVMLWAPIVYILLARLIDGFGIGLAVTLVPLY 120
QY 176 ISEIAPSEIR-GLLNTLPQPSG-SGGMFLSYCMVFGMSLSPSPDWRIMLGVLAIPSLFFF 233
Db 121 ISETAPHRXSWGXXNTLPQFIGVXGGMFLSYCMVFGMSLMPKPDWRMLMGVLSIPSLXYF 180
QY 234 GLTIFYLPESPRWLVSCKGRMAEAKKVLQKLRGKDDVSGELSLLLEGLVGGDTSIEEYII 293
Db 181 GLTVFYLPESPRWLVSCKGRMAEAKRVXQRLRGREDVSXEXALLVEGLGVGKDTRIXEYII 240
QY 294 GPATEAADLVTDDKEQITLYGPEEGQSWIARPSKGPIMLGSVLSASRRHGMVNQSV 353
Db 241 GPATEAADLVTDDKEQITLYGPEEGQSWIARPSKGPIMLGSVLSASRRHGMVNQSV 300
QY 354 LMDPIVTLFGSVHENMPQAGGMRSTLFPNFGSMFVTDQHAKNEQWDEENLHRDDEEYA 413
Db 301 LMDPIVTLFGSVHENMPQAGGMRSTLFPNFGSMFVTDQHAKNEQWDEENLHRDDEEYA 360
QY 414 SDGAGGDYEDNLHSPLLSRQATGAEGKDIVHHGHRGSALSMMRQTLLEGGDGVSSTDIG 473
Db 361 SDGAGGDYEDNLHSPLLSRQATGAEGKDIVHHGHRGSALSMMRQTLLEGGDGVSSTDIG 420
QY 474 GGWQLAWKWEKEGNGRKEGGKRVYLHQEGVPGRRGSIIVSLPGGDDVFEGSEFVHAA 533
Db 421 GGWQLAWKWEKEGNGRKEGGKRVYLHQEGVPGRRGSIIVSLPGGDDVFEGSEFVHAA 480
QY 534 ALVSQSALFSKGLAEPRMSDAAMVHPSEVAAGSRWKDLFEPGVRALLVGVIQILOQF 593
Db 481 ALVSQSALFSKGLAEPRMSDAAMVHPSEVAAGSRWKDLFEPGVRALLVGVIQILOQF 540
QY 594 AGINGVLYYTPQILEQAGVAVILSKFGLSSASASILISSLTTLMLPCIGFAMLLMDLSG 653
Db 541 AGINGVLYYTPQILEQAGVAVILSKFGLSSASASILISSLTTLMLPCIGFAMLLMDLSG 600
QY 654 RRFLLLGTIPILIASLVILVSNLIDGLTLAHALLSTVSIVVYFCFWMGFGPIPNILCA 713
Db 601 RRFLLLGTIPILIASLVILVSNLIDGLTLAHALLSTVSIVVYFCFWMGFGPIPNILCA 660
QY 714 EIFPTRVRGLCIAICAFTEFWIGDIIIVTYSLPVMLNAIGLAGVFSIYAVVCLISFVFVLK 773
Db 661 EIFPTRVRGLCIAICAFTEFWIGDIIIVTYSLPVMLNAIGLAGVFSIYAVVCLISFVFVLK 720
QY 774 VPETKGMPLVITEFFAVGAKQAAKA 800
Db 721 VPETKGMPLVITEFFAVGAKQAAKA 747

RESULT 5
ADG47906

ID ADG47906 standard; protein; 747 AA.
XX

AC ADG47906;
XX

DT 11-MAR-2004 (first entry)
XX

DE Corn Arabidopsis-like sugar transport protein #1.
XX

KW Arabidopsis-like sugar transport protein;
KW Beta-vulgaris-like sugar transport protein; transgenic; physical mapping;
KW corn; plant.
XX

OS Zea mays.
XX

FH Key Location/Qualifiers
FT Misc-difference 129

FT /note= "Encoded by NAT"
FT Misc-difference 133
FT /note= "Encoded by TGN"
FT Misc-difference 134
FT /note= "Encoded by TNG"
FT Misc-difference 144
FT /note= "Encoded by AGN"
FT Misc-difference 178
FT /note= "Encoded by ATN"
FT Misc-difference 195
FT /note= "Encoded by GTN"
FT Misc-difference 207
FT /note= "Encoded by NTG"
FT Misc-difference 218
FT /note= "Encoded by NGG"
FT Misc-difference 220
FT /note= "Encoded by ANG"
FT Misc-difference 236
FT /note= "Encoded by TNA"
XX
PN US2002199217-A1.
XX
XX 26-DEC-2002.
XX
XX 17-JAN-2002; 2002US-00051909.
XX
XX 24-APR-1998; 98US-0083044P.
XX 14-APR-1999; 99US-00291922.
XX
XX (HELE/) HELENTJARIS T G.
XX Helentjaris TG;
XX WPI; 2004-040967/04.
XX N-PSDB; ADG47905.
XX
XX New isolated polynucleotide encoding a polypeptide having sugar transport
PT protein activity, for producing a transformed plant and for use as probes
PT in physical mapping.
XX
XX Example 3; SEQ ID NO 2; 71pp; English.
XX
XX The invention relates to Arabidopsis-like or Beta-vulgaris-like sugar
CC transport proteins and their corresponding nucleic acid sequences. The
CC sequences of the invention are useful to transform a cell. These are also
CC useful to produce a transgenic plant. Probes derived from sequences
CC encoding sugar transport protein may be used for physical mapping. The
CC present sequence is corn Arabidopsis-like sugar transport protein.
XX
XX Sequence 747 AA;
SQ
Query Match 86.3%; Score 3517; DB 8; Length 747;
Best Local Similarity 91.6%; Pred. No. 0;
Matches 684; Conservative 27; Mismatches 34; Indels 2; Gaps 2;
Qy 56 MSGAVLVAIVASIGNLLQGWDNATIAAAVLYIKKEFQIQNEPTVEGLIVSMSLIGATIVT 115
Db 1 MGGAVMVAIAAASIGNLLQGWDNATAGAVLYIKKEFNLOSEPLIEGLIVAMFLIGATVIT 60
Qy 116 TFSGLSDSIGRRPMLILSSILYFFSGLIMLWSPNVYVLLARFVDGFGIGLAVTLVPLY 175
Db 61 TSPGPRADCVGRRPMLVASAVLYFVSGLVMLWAPIVYILLARLIDGFGIGLAVTLVPLY 120
Qy 176 ISEIAPSEIR-GLLNTLPQFSG-SCGMFLSYCMVFGMSLSPSPDWRIMGLVALPSLFFF 233
Db 121 ISETAPHRXSWGXXNTLPQFVIGXGGMFLSYCMVFGMSLMPKPDWRMLGLVLSIPSLXYF 180
Qy 234 GLTIFYLPESPRWLVSKGMAEAKVQLRKGKDDVSGELSLLEGLVGGDTSEIEYII 293
Db 181 GLTVFYLPESPRWLVSKGMAEAKRVXQRLRGREDVSEXALLVEGLVGKDTRIXEYII 240
Qy 294 GPATEAADDLVTDGKEQITLYGPPEGQSWIARPSKPIMLGSVLSLASRHGSMVNSQVSP 353

Db 241 GPATEAADDLVTDGKEQITLYGPPEGQSWIARPSKPIMLGSVLSLASRHGSMVNSQVSP 300
Qy 354 LMDPIVTLFGSVHENMPOAGGSMRSTLFPNFGSMFVSVDQHAQNEQWDEENLHRDDEEYA 413
Db 301 LMDPIVTLFGSVHENMPOAGGSMRSTLFPNFGSMFVSVDQHAQNEQWDEENLHRDDEEYA 360
Qy 414 SDGAGGDYEDNLHSPLLSRQATGAEGKDIVHHGHRGSALSMMRQTLLGEGGDGVSSTDIG 473
Db 361 SDGAGGDYEDNLHSPLLSRQATGAEGKDIVHHGHRGSALSMMRQSLLEGEGGDGVSSTDIG 420
Qy 474 GGWQLAWKWEKEGNGRKEGGFKRVYLHQEGVPGSRGSIIVSLPGGGDVFESEFVHAA 533
Db 421 GGWQLAWKWEKEGNGRKEGGFKRVYLHQEGVPGSRGSIIVSLPGGGDVFESEFVHAA 480
Qy 534 ALVQSQALFSKGLAEPRMSDAAMVHPSEVAAGKSRWKDLFEPGVRALLVGVGIQILQQF 593
Db 481 ALVQSQALFSKGLAEPRMSDAAMVHPSEVAAGKSRWKDLFEPGVRALLVGVGIQILQQF 540
Qy 594 AGINGVLYYTPQILEQAGVAVILSKFGLSSASASILISSLTLLMLPCIGFAMLLMDLSG 653
Db 541 AGINGVLYYTPQILEQAGVAVILSKFGLSSASASILISSLTLLMLPCIGFAMLLMDLSG 600
Qy 654 RRFLLLGTIPILIASLVILVNSNLDLGLTLAALLSTVSVIVVFFCCFVMGFGPIPNILCA 713
Db 601 RRFLLLGTIPILIASLVILVNSNLDLGLTLAALLSTVSVIVVFFCCFVMGFGPIPNILCA 660
Qy 714 EIFPTRVRGLCIAICAFTEFWIGDIIIVTYSLPVMLNAIGLAGVFSIYAVVCLISFVVELK 773
Db 661 EIFPTRVRGLCIAICAFTEFWIGDIIIVTYSLPVMLNAIGLAGVFSIYAVVCLISFVVELK 720
Qy 774 VPETKGMPLLEVITEFFFAVGAQAQAAKA 800
Db 721 VPETKGMPLLEVITEFFFAVGAQAQAAKA 747
RESULT 6
ADG47941
ID ADG47941 standard; protein; 740 AA.
XX
AC ADG47941;
XX
DT 11-MAR-2004 (first entry)
XX
DE Arabidopsis thaliana-like sugar transport protein #2.
XX
XX Arabidopsis-like sugar transport protein;
KW Beta-vulgaris-like sugar transport protein; transgenic; physical mapping;
KW mouse-ear cress; plant.
XX
OS Arabidopsis thaliana.
XX
PN US2002199217-A1.
XX
PD 26-DEC-2002.
XX
PF 17-JAN-2002; 2002US-00051909.
XX
XX 24-APR-1998; 98US-0083044P.
PR 14-APR-1999; 99US-00291922.
XX
XX (HELE/) HELENTJARIS T G.
XX Helentjaris TG;
XX WPI; 2004-040967/04.
XX
XX New isolated polynucleotide encoding a polypeptide having sugar transport
PT protein activity, for producing a transformed plant and for use as probes
PT in physical mapping.
XX
XX Example 3; SEQ ID NO 37; 71pp; English.
XX
XX The invention relates to Arabidopsis-like or Beta-vulgaris-like sugar

DE Corn Arabidopsis-like sugar transport protein #2.
XX Arabidopsis-like sugar transport protein;
KW Beta-vulgaris-like sugar transport protein; transgenic; physical mapping;
KW corn; plant.
XX
OS Zea mays.
XX US2002199217-A1.
XX 26-DEC-2002.
XX 17-JAN-2002; 2002US-00051909.
XX 24-APR-1998; 98US-0083044P.
XX 14-APR-1999; 99US-00291922.
XX (HELE/) HELENTJARIS T G.
XX Helentjaris TG;
PI WPI; 2004-040967/04.
XX N-PSDB; ADG47911.
DR
XX
PT New isolated polynucleotide encoding a polypeptide having sugar transport
PT protein activity, for producing a transformed plant and for use as probes
PT in physical mapping.
XX
XX Example 3; SEQ ID NO 8; 71pp; English.
PS
XX The invention relates to Arabidopsis-like or Beta-vulgaris-like sugar
CC transport proteins and their corresponding nucleic acid sequences. The
CC sequences of the invention are useful to transform a cell. These are also
CC useful to produce a transgenic plant. Probes derived from sequences
CC encoding sugar transport protein may be used for physical mapping. The
CC present sequence is corn Arabidopsis-like sugar transport protein.
XX
SQ Sequence 737 AA;

Query Match 65.6%; Score 2674; DB 8; Length 737;
Best Local Similarity 69.9%; Pred. No. 3.8e-250;
Matches 522; Conservative 92; Mismatches 115; Indels 18; Gaps 8;

QY 56 MSGAVLVAIVASIGNLLQGDNDATIAAAVLIYIKKEFQLQNEPTVEGLIVSMSLICATIVT 115
Db 1 MKGAVLVAIAASIGNFLQGDNDATIAAGANGYIKKDLALGT--TMERLVVGMSLICATVIT 58

QY 116 TFGPLSDSIGRRPMLILSSILYFFSGLIMLWSPNVYVLLARFVDGFGIGLAVTLVPVLY 175
Db 59 TCSGPIADWLGRRRPMMIISVLYFLGGLVMLWSPNVYVLLARLLDGFGLAVTLVPVY 118

QY 176 ISEIAPSEIRGLLNTLPQFSGGGMFLSYCMVFGMSLSPSPDRIMLGVLAIPLSFFGL 235
Db 119 ISETAPSEIRGSLNLTLPQFSGGGMFLSYCMVFGMSLSPSPDRIMLGVLAIPLSFFGL 178

QY 236 TIFLPSPRWLVSKGRMAEAKVQLKRGKDDVSGELSLLEGLVGGDTSIIEYIIGP 295
Db 179 TIFLPSPRWLVSKGRMLEAKVQLKRGKDDVSGELSLLEGLVGGDTSIIEYIIGP 238

QY 296 ATEAADLVTDGDKQITLYGPEGQSWIARPSKGPIMLGSVLSLASRHGSMVNSQSVPLM 355
Db 239 ADDVADGHEHATEKDRIKLYSQAGLSWLSKPVGTQSSIG---LASHHGSINQSMPLM 294

QY 356 DPLVTLFGSVHENMPQAG--GSMRSTLFPNFGSMFSTVDQAKNEQWDEENLHRDDEYA 413
Db 295 DPLVTLFGSIHEKLPETGARGSMRSTLFPNFGSMFSTVDQAKNEQWDEESLQREDYD 354

QY 414 SDGAGDYEDNLHSPLLSRQATGAEGKDIVHGHGRGSAL--SMRQTLLEG--GDGVSSTD 471
Db 355 SDATRGSDDDLHSPILSRQITSLK--KDLPPPPSHGSLGSMRRHSLMQSGEGGSGTG 413

QY 472 IGGWQLAWKWEKEGNGRKEGGFKRVYLHQGVPGSRRGSIIVSLPGGDFEGSEFVH 531
XX

Db 414 IGGWQLAWKWKD-KEDGKQGGFKRIYLHEEGVSASRRGSIVSIPGEG-----EFVQ 466
QY 532 AAALVSQSALFSKGLAEPMSDAAMVHPSEVAAGSRWKDLFEFGVRRALLVGVGIQILQ 591
Db 467 AAALVSQPALYSKELIDGHPVGPAMVHPSETASKGSPWKALLEPGVKHALVGVGIQILQ 526
QY 592 QFAGINGVLYYTPQILEQAGVAVILSKFGLSSASASILISSLTLLMLPCIGFAMLLMDL 651
Db 527 QFSGINGVLYYTPQILEEAGVEVLLSDIGIGSESASFLLISAFITFLMLPCIGVAMKMDV 586
QY 652 SGRFLLLTGTPILIASLVILVNSNLDLGLTALHALLSTVSVIVVFCFVMGFGPIPNIL 711
Db 587 SGRQLLLTTTPVLIVSLIILVIGSLVNFNGVAHAIAISTVCVVVYFCCFVMGFGPIPNIL 646
QY 712 CAEIFPTRVRGLCIAICAFTEWIGDIIIVTYSLPVMLNAIGLAGVFSIYAVVCLISFVVF 771
Db 647 CSEIFPTRVRGLCIAICAFTEWIGDIIIVTYSLPVMLGSLGLGVFAIYAVVCFISWIFVF 706
QY 772 LKVPETKGMPLVITEFFAVGAKQAAA 798
Db 707 LKVPETKGMPLVISEFFSVGAKQAAAS 733

RESULT 10
ABR93252
ID ABR93252 standard; protein; 729 AA.
XX
AC ABR93252;
XX
DT 31-MAY-2002 (first entry)
XX
DE Herbicidally active polypeptide SEQ ID NO 2463.
XX
KW Herbicidal; plant; agriculture; herbicide.
XX
OS Arabidopsis thaliana.
XX
PN WO200210210-A2.
XX
PD 07-FEB-2002.
XX
PF 28-AUG-2001; 2001WO-EP009892.
XX
PR 28-AUG-2001; 2001WO-EP009892.
XX
PA (FARB) BAYER AG.
XX
PI Tietjen K, Weidler M;
XX
DR WPI; 2002-269010/31.
XX
PT Identifying plant target proteins for herbicidally active compounds,
PT comprising aligning and comparing nucleic acid or amino acid sequences
PT from plant with nucleic acid or amino acid sequences from non-plant
PT organisms.
XX
PS Claim 5; SEQ ID NO 2463; 261pp + Sequence Listing; English.
XX
CC The invention relates to identifying target proteins (ABR90790-ABR94016)
CC for herbicidally active compounds, comprising aligning and comparing
CC nucleic acid or amino acid sequences from plant with nucleic acid or
CC amino acid sequences from non-plant organisms using suitable search
CC parameters, where plant sequences having an E-value greater by a factor
CC of 3 than the E-value of most similar non-plant sequences are selected.
CC The polypeptides or nucleic acids encoding them are useful for
CC identifying modulators. The identified modulators are useful as
CC herbicides
XX
SQ Sequence 729 AA;

Query Match 63.4%; Score 2584; DB 5; Length 729;
Best Local Similarity 68.8%; Pred. No. 2.1e-241;
Matches 512; Conservative 89; Mismatches 123; Indels 20; Gaps 8;

Db 582 RTLLTTPILIASLLVLVISNLVHMNSIVHAVLSTVSWLYFCFFVMGFGPAPNILCSE 641
Qy 715 IPFTRVRGLCIAICAFTHWIGDIIIVTYSLPVMLNAIGLAGVFSIYAVVCLISFVFPLKV 774
Db 642 IPFTRVRGICIAICALTFWICDIIIVTYSLPVLLKSIAGVFGMYAIVCCISWVFFIKV 701
Qy 775 PETKGMPLVITEFFAVGAKQAAA 798
Db 702 PETKGMPLVITEFFSVGARQAEA 725

RESULT 12
ABB91078
ID ABB91078 standard; protein; 734 AA.
AC ABB91078;
XX
DT 31-MAY-2002 (first entry)
XX
DE Herbicidally active polypeptide SEQ ID NO 289.
XX
KW Herbicidal; plant; agriculture; herbicide.
XX
OS Arabidopsis thaliana.
XX
PN WO200210210-A2.
XX
PD 07-FEB-2002.
XX
PF 28-AUG-2001; 2001WO-EP009892.
XX
PR 28-AUG-2001; 2001WO-EP009892.
XX

PA (FARB) BAYER AG.
XX
PI Tietjen K, Weidler M;
XX
WPI; 2002-269010/31.
XX
Identifying plant target proteins for herbicidally active compounds,
PT comprising aligning and comparing nucleic acid or amino acid sequences
PT from plant with nucleic acid or amino acid sequences from non-plant
PT organisms.
XX
Claim 5; SEQ ID NO 289; 261pp + Sequence Listing; English.
PS
XX
The invention relates to identifying target proteins (ABB90790-ABB94016)
CC for herbicidally active compounds, comprising aligning and comparing
CC nucleic acid or amino acid sequences from plant with nucleic acid or
CC amino acid sequences from non-plant organisms using suitable search
CC parameters, where plant sequences having an E-value greater by a factor
CC of 3 than the E-value of most similar non-plant sequences are selected.
CC The polypeptides or nucleic acids encoding them are useful for
CC identifying modulators. The identified modulators are useful as
CC herbicides
XX
SQ Sequence 734 AA;

Query Match 55.8%; Score 2275; DB 5; Length 734;
Best Local Similarity 61.2%; Pred. No. 2.5e-211;
Matches 462; Conservative 111; Mismatches 146; Indels 36; Gaps 14;
Qy 56 MSGAVLVAIVASIGNLLQWDNATIAAAVLYIKKEFQLQNEPT-VEGLIVSMSLIGATIV 114
Db 1 MKGATLVALAATIGNFLQWDNATIAGAMVYINKDLNL---PTSVQGLVAVMSLIGATIVI 57
Qy 115 TTFSGPLSDSICRRPMLILSSILYFFSGLIMLWSPNVYVLLARFVDFGIGLAVTLVPL 174
Db 58 TTCSGPISDWLGRPRMLILSSVMYFVCGLIMLWSPNVYVLCFARLLNGFAGLAVTLVPV 117
Qy 175 YISEIAPSEIRGLINTLPQFSGGGMFLSYCMVFGMSLSPSPDWRIMLVLAIPSLFFFG 234

Db 118 YISETAPPEIRQLNTLPQFLGSGGMFLSYCMVFTMSLSDSPSWRAMLGVLSPISLLYLF 177
Qy 235 LTIFYLPESPRWLVSGRMAEAKVLQKLRGKDDVSGELSLLEGLVCGDTSIEEYIIG 294
Db 178 LTIFYLPESPRWLVSGRMDEAKRVLQQLCGREDVTDEMALLVEGLDIGEKTMEDLLVT 237
Qy 295 PATEAADDLVTDGKE-QITLYGPEEGQSWIARPSKGPIMLGSLVSLASRHGSMVNSQV 353
Db 238 LEDHEGDDTLETVDGQMLRYGTHENQSYLARVPVEQ---NSSLGRSRHGSANQSMI 294
Qy 354 LMDPIVTLFGSVHENMPQAGGSMRSTLFPNFGSMFSVTDQ--HAKNEQWD---EENLHRD 408
Db 295 LKDPVLNLFGLSHEKMPAEGNTRSGIFPHFGSMFSTTADAPHGKPAHWEKDIESHYNKD 354
Qy 409 DEEYASDGAGGDYED--NLHSPLLSRQATGAEGKDIVHHGHRGSALSMRRQTLLGEGGD 465
Db 355 NDDYATDDGAGDDDDSDNDLRSPLMSRQTTSMD-KDMIHPHTSGSTLSMRRHSTLMQ-GN 412
Qy 466 GVSSTDIGGWQLAWKWEKEGNGRKEGPKRVYLHQEGVPGSRRGSIVSLPGGVDVFE 525
Db 413 GESSMGIGGGWHMGYRYENDE-----YKRYLKEDEGAE-SRRGSIISIPGGPD--G 460
Qy 526 GSEFVHAAALVSQSALFSKGLAEPRMSDAMVHPSEVAAGSRWKDLFEPGVRRALLVGV 585
Db 461 GGSYIHASALVSRSVLGPKS-----VHGSAMVPEKIAASGPLWSALLEPGVKRALVGV 515
Qy 586 GIQLQQFAGINGVLYTPQILEQAGVAVILSKFGLSSASASILISSLTLLMLPCIGFA 645
Db 516 GIQLQQFSGINGVLYTPQILERAGVDILLSSGLSSISASFLISGLTLLMLPAIIVA 575
Qy 646 MLLMDLSGRRFLLGTIPILIASLVILVSNLIDLTGLAHALLSTVSVIVYFCFVMGFG 705
Db 576 MRLMDVSGRRSLLLTWIPVLIVSLVVLVISELIHISKVNAALSTGCVVLYFCFFVMGYG 635
Qy 706 PIPNILCAEIPFTRVRGLCIAICAFTHWIGDIIIVTYSLPVMLNAIGLAGVFSIYAVVCLI 765
Db 636 PIPNILCSEIFPTRVRGLCIAICAMVFWIGDIIIVTYSLPVLLSSIGLVGVFSIYAAVCVI 695
Qy 766 SFVFLKVPETKGMPLVITEFFAVGAKQAAA 800
Db 696 SWIFVYMKVPETKGMPLVITEFFAVGAKQAAA 729

RESULT 13
ABB92622
ID ABB92622 standard; protein; 729 AA.
XX
AC ABB92622;
XX
DT 31-MAY-2002 (first entry)
XX
DE Herbicidally active polypeptide SEQ ID NO 1833.
XX
KW Herbicidal; plant; agriculture; herbicide.
XX
OS Arabidopsis thaliana.
XX
PN WO200210210-A2.
XX
PD 07-FEB-2002.
XX
PF 28-AUG-2001; 2001WO-EP009892.
XX
PR 28-AUG-2001; 2001WO-EP009892.
XX
PA (FARB) BAYER AG.
XX
PI Tietjen K, Weidler M;
XX
DR WPI; 2002-269010/31.
XX
PT Identifying plant target proteins for herbicidally active compounds,
PT comprising aligning and comparing nucleic acid or amino acid sequences

PT from plant with nucleic acid or amino acid sequences from non-plant
PT organisms.
PS Claim 5; SEQ ID NO 1833; 261pp + Sequence Listing; English.
XX
CC The invention relates to identifying target proteins (ABB90790-ABB94016)
CC for herbicidally active compounds, comprising aligning and comparing
CC nucleic acid or amino acid sequences from plant with nucleic acid or
CC amino acid sequences from non-plant organisms using suitable search
CC parameters, where plant sequences having an E-value greater by a factor
CC of 3 than the E-value of most similar non-plant sequences are selected.
CC The polypeptides or nucleic acids encoding them are useful for
CC identifying modulators. The identified modulators are useful as
CC herbicides
XX
SQ Sequence 729 AA;

Query Match 55.1%; Score 2245; DB 5; Length 729;
Best Local Similarity 60.4%; Pred. No. 2.1e-208;
Matches 457; Conservative 110; Mismatches 144; Indels 46; Gaps 14;

Qy 56 MSGAVLVAIVASIGNLQWDNATIAAAVLYIKKEFQLONEPTVEGLIVMSLIGATIVT 115
Db 1 MRSVVLVALAAIGNMLQWDNATIAAGAVIYIKKEFHLEKEPKIEGLIVAMSLIGATLIT 60

Qy 116 TFGSPLSDSIGRRPMLILSSILYFFSGLIMLWSPNVYVLLARFVDGFGIGLAATLVPLY 175
Db 61 TFGSPVSDKVGRRSMILSSVLYFLSSIVFWSPNVYVLLFARLLDGGFGLAATLVPIY 120

Qy 176 ISEIAPSEIRGLNTLPQSGSGMFLSYCMVFGMSLSPSPDWIRIMLVLAIPSLFFGL 235
Db 121 ISETAPSEIRGLNTLPQSGSGMFLSYCLVFGMSLQSPSWRLMLGVLSIPSIAFVL 180

Qy 236 TIFYLPESPRWLVSKGRMAAKVQLKRGKDDVSGELSLLEGLVGGDTSEIYYIGP 295
Db 181 AAPFLPESPRWLVSKGRMDEARQVQLRGREDVSGELALLVEGLVGKDTSEIYYIGP 240

Qy 296 ATBAADDLVTGDKEQITLYGPEEGQSWIARPSKGPIMLGSVLSASRHGSMVNQSVPLM 355
Db 241 DNEENEGGNELPRKQIKLYGPEQGQSWMAKPVKQ---SSLASARQSGMLPRGGSML 296

Qy 356 DPIVTLFGSVHENMP--QAGSSMRSTLFPNFGSMFVTDQHAQNEQWDEENLHRDDEYA 413
Db 297 DPLVTLFGSIHENLPSENNMASSRMLFPNMGSILGMGR--QESQWDPE---RNED-- 349

Qy 414 SDGAGGDYEDNLHSPLLSRQATGAEGKDIVHHGHRGSALSMMRQTL-LGEGGDVSSTDI 472
Db 350 -----SSDQDENLNSPLLSPQT--EPDD--YHQRTVGTMRHQSSLFMANVGETATATSI 401

Qy 473 GGGWQLAWKWEKEGNGRK-EGGFKRVYLHQE-----GVPGSRGSIIVSL-PGGGDV 523
Db 402 GGGWQLAWKYNKVGADGKRVNGGLQRMVYIHEETANNNTNNIPFSRRGSLLSFHEPGDGH 461

Qy 524 FEGSEFVHAAALVSQSALFSGLABPRMSDAAVHPSEVAAGSRWKDLFEPGVRRALLY 583
Db 462 DVNGYVQAAALVSQASMPGKGETAML-----PKEV-KDGPGRWRELKEPGVXRALMV 514

Qy 584 GVGIQLIQFAGINGVLYYTPQILEQAGVAVILSKFGLSSASASILISSLTLLMLPCIG 643
Db 515 GVGLQLIQFAGINGVMYYTPQILEETGVSSLLTNLGISAESASLLISALTLLMLPCI- 573

Qy 644 FAMLMDLSGRRFLLLTGTPILIASLVILVSNLIDGLTAHALLSVSVIVYFCCFVMG 703
Db 574 -----LVSMRSLMLSTIPILILSLVTLVIGSLVNLGGSINALISTASVTVYLSCFVMG 626

Qy 704 FGPIPNILCAEIFFPTRVRGLCIACAFTFWIGDIIIVTYSLPWMLNAIGLAGVFSIYAVVC 763
Db 627 FGAIPNLCSEIFPTSVRGLCITICALTFWICDIIIVTYSLPWMLKSIAGVFGIYAVVC 686

Qy 764 LISFVFVLKVPETKGMPLVETFEFFAVGAKQAAKA 800
Db 687 AVAVVFVLKVPETKGMPLVETFEFFSVGAKQDAAA 723

RESULT 14
AAU97205
ID AAU97205 standard; protein; 486 AA.
XX
AC AAU97205;
XX
DT 27-AUG-2002 (first entry)
XX
DE Portion of a soybean sugar transport protein encoded by ssl.pk0022.fl.
XX Soybean; Arabidopsis thaliana-like sugar transport protein;
KW carbohydrate transport; grain filling; annual field crop; plant.
XX
OS Glycine max.
XX
PN US6383776-B1.
XX
PD 07-MAY-2002.
XX
PF 14-APR-1999; 99US-00291922.
XX
PR 24-APR-1998; 98US-0083044P.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;
XX
DR WPI; 2002-453364/48.
DR N-PSDB; ABK51966.
XX
PT New nucleic acid encoding plant sugar-transport proteins, useful for
PT preparing transgenic plants with altered carbohydrate distribution.
XX
PS Claim 1; Fig 1; 54pp; English.
XX
CC The present invention relates to the isolation of plant polynucleotide
CC sequences encoding an Arabidopsis thaliana-like sugar transport protein
CC or Beta vulgaris-like sugar transport protein. The polynucleotide
CC sequences are useful for altering the level of sugar transport proteins
CC in plants, i.e. for control of carbohydrate transport and distribution in
CC plant cells, e.g. during grain filling of annual field crops (e.g. corn,
CC rice, soybeans, and wheat), and, for studying carbohydrate flows and
CC sugar transport. The polynucleotide sequences can also be used to isolate
CC cDNA sequences and genes that encode homologues of the new proteins. The
CC present sequence represents a portion of a soybean Arabidopsis thaliana-
CC like sugar transport protein
XX
SQ Sequence 486 AA;

Query Match 41.0%; Score 1670.5; DB 5; Length 486;
Best Local Similarity 68.1%; Pred. No. 8.1e-153;
Matches 333; Conservative 57; Mismatches 84; Indels 15; Gaps 8;

Qy 308 DKEQITLYGPEEGQSWIARPSKGPIMLGSVLSASRHGSMVNQSVPLMDPIVTLFGSVHE 367
Db 5 ERDQIKLYGPEEGQSWVARPVAGPNSVG---LVSRRKGSMAFPS-SLVDPLVTLFGSVHE 59

Qy 368 NMPQAGGSMRSTLFPNFGSMFVTDQHAQNEQWDEENLHRDDEYASDGAGGDYEDNLHS 427
Db 60 KLPETG---STLFPFGSMFVGGNQPRNEDWDEESLAREGDDYVSD--AGSDDDNLQS 113

Qy 428 PLLSRQATGAEGKDIVHHGHRGSALSMMRQTL-LGEGGDVSSSTDIGGGWQLAWKWEKE 486
Db 114 PLISRQTTSLD-KDIPPHAHNSLA-SMRQGSLLHNSGEPSTGTGIGGGWQLAWKWSERE 171

Qy 487 GENGKKEGGFKRVLHQEGVPGSRGSIIVSLPGGGDVFESEFVHAAALVSQSALFSKGL 546
Db 172 GPDGKKEGGFKRIYLHQDGGSGSRGSGVSLP-GGDLPTDSEVVQAAALVSPALYNEDL 230

Qy 547 AEPRMSDAAVHPSEVAAGSRWKDLFEPGVRRALLYGVGIQLIQFAGINGVLYYTPQI 606
Db 231 MRQRPVGPAMTHPSETIAKGPSWSDLFEPGVKHALIVGVGMQILQGFSGINGVLYYTPQI 290

QY	607	LEQAGVAVILSKFGLSSASASILISSLTTLMLPCIGFAMLLMDLSGRRFLLLGTIPILI	666
Dd	291	LEQAGVGYLISLGLGSTSSSFLISAFTLLMLPCIAIAMRLMDISGRRTLLSLSTIPVLI	350
QY	667	ASLVILVSNNLDLGTLAHALLSTVSIVVFCCFVMGFGPIPNILCAEIFPTRVRGLCIA	726
Dd	351	AALLILVLGSLVDLGSTANASISTISIVVFCCFFVMGFGPIPNILCAEIFPTRVRGLCIA	410
QY	727	ICAFTEFWICDIIVTYSLPVMLNAIAGLVFSIYAVVCCLISFVFVKVPETKGMPLLEVIT	786
Dd	411	ICALTFWICDIIVTYTLPVMLNSVGLAGVFGIYAVVCFIAWFVFLKVPETKGMPLVII	470
QY	787	EFFAVGAKQ	795
Dd	471	EFFSVGAKQ	479

RESULT 15
ABU08330
ID ABU08330 standard; protein; 486 AA.

DT 29-MAY-2003 (first entry)

DE Soybean sugar transport protein #2.

KW Arabidopsis thaliana-like sugar transport protein; corn; rice; wheat;
KW plant sugar transport protein; carbohydrate transport; soybean;
KW carbohydrate distribution; plant.

Glycine max.

PN US2002178468-A1.

PD 28-NOV-2002.

PF 17-JAN-2002; 2002US-00051902.

PR 24-APR-1998; 98US-0083044P.

PR 14-APR-1999; 99US-00291922.

PA (ALLE/) ALLEN S M.

PA (HITZ/) HITZ W D.

PA (KINN/) KINNEY A J.

PA (TING/) TINGEY S V.
YY

PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;

DR WPI; 2003-340957/32.

DR N-PSDB; ABX93202.

PT Novel plant sugar transport proteins and nucleic acid encoding the
PT protein useful for producing transgenic plants having altered levels of
PT sugar transport protein.

PS Claim 5; Fig 1; 56pp; English.

The present invention relates to the isolation of Arabidopsis thaliana-like or Beta vulgaris-like sugar transport proteins, and the polynucleotide sequences encoding them. The plant sugar transport proteins of the invention have been isolated from corn, rice, soybean, and wheat. The polypeptides of the invention may be used for altering the level of expression of a sugar transport protein in a host cell, by transforming a host cell with a chimeric construct encoding all, or a portion of the sugar transport protein, in sense or antisense orientation. Particularly, the polypeptides may provide a means to control carbohydrate transport and distribution in plants. ABU08326-ABU08333 represent Arabidopsis thaliana-like sugar transport proteins

Sequence 486 AA;

Query Match	41.0%;	Score 1670.5;	DB 6;	Length 486;
Best Local Similarity	68.1%;	Pred. No. 8.1e-153;		
Matches 333;	Conservative 57;	Mismatches 84;	Indels 15;	Gaps 8;
QY	308	DKEQITLYGPEEGQWIARPSKGPIMLGSVLSLASHGSMVNQSVPLMDPIVTLFGSVHE	367	
Db	5	EKDQIKLYGPEQGQSWARPVAGPNSVG----	LVSRRKSGMANPS-SLVDPLVTLFGSVHE	59
QY	368	NMPQAGGSMRSTLFPNFGSMFVTDQAHAKNEQWDEENLHRDDEEYASDAGAGDYEDNLHS	427	
Db	60	KLPETG----STLFPHFGSMFVSGNQPRNEDWDEESLAREGDDYVSD--	AGDSDDNLQS	113
QY	428	PLLSRQATGAEGKDIVHHGHRGSALSMMRQTLL-GEGGDGVSTDIGGGWQLAWKWEKE	486	
Db	114	PLISRQTTSLD-KDIPPHAHSNLA-SMRQGSLLHSGSEPTGSTGIGGGWQLAWKWSERE	171	
QY	487	GENGRKEGGFKRYLHQEGVPGSRRGSIIVSLPGGGDVFECSSEFVHAAALVSQSALFSKGL	546	
Db	172	GPDGKKEGGFKRIYLHQDGGSGRRGVSVSLP-GGDLPTDSEVVQAAALVSQPALYNEDL	230	
QY	547	AEPRMSDAAMVHPSEVAAKGRWKDLFEPGVRRALLVGVGIQILQQFAGINGVLYYTPQI	606	
Db	231	MRQRPVGPAMIHPSETIAKGPSWSDLFEPGVKHALIVGVGMQILQQFSGINGVLYYTPQI	290	
QY	607	LEQAGVAVILSKFGLSSASASILSSLTTLMLPCIGFAMLLMDLSGRRFLLGTIPII	666	
Db	291	LEQAGVGYLLSSLGLGSTSSSFLISAVTTLLMLPCIAIAMRLMDISGRRTLLSTIPVLI	350	
QY	667	ASLVILVSNLIDLGTIAHALLSTVSVIVYFCFVMGFGPIPNILCAEIFFTRVRGLCIA	726	
Db	351	AALLILVLGSLVDLGSTANASISTISVIVYFCFVMGFGPIPNILCAEIFFTRVRGLCIA	410	
QY	727	ICAFTFWIGDIIVTYSLPVMNLNAIGLAGVFSIYAVVCLISFVFVFLKVPETKGMPELVIT	786	
Db	411	ICALTFWICDIIVTYTLPVMLNSVGLAGVFGIYAVVCFIAWVFVFLKVPETKGMPELVII	470	
QY	787	EFFAVGAKQ	795	
Db	471	EFFSVGAKQ	479	

Search completed: October 13, 2004, 11:38:26
Job time : 85.4139 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2004, 11:33:23 ; Search time 24.8337 Seconds
(without alignments)
2136.387 Million cell updates/sec

Title: US-10-051-909-32
Perfect score: 4075
Sequence: 1 IRSGSWLAVQTFPFDPDLRR.....PLEVITEFFAVGAKQAQAAKA 800

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3517	86.3	747	3	US-09-291-922-2
2	2674	65.6	737	3	US-09-291-922-8
3	2584	63.4	729	3	US-09-291-922-29
4	1670.5	41.0	486	3	US-09-291-922-10
5	1424	34.9	345	3	US-09-291-922-14
6	943	23.1	228	3	US-09-291-922-16
7	562	13.8	131	3	US-09-291-922-6
8	548	13.4	488	4	US-10-162-012-46
9	496	12.2	523	3	US-09-291-922-24
10	489.5	12.0	513	3	US-09-291-922-20
11	489	12.0	529	3	US-09-291-922-28
12	488	12.0	549	3	US-09-291-922-30
13	485	11.9	510	3	US-09-291-922-22
14	463.5	11.4	539	3	US-09-291-922-26
15	447	11.0	476	4	US-09-489-039A-11933
16	440.5	10.8	514	4	US-09-489-039A-11902
17	427	10.5	501	4	US-09-489-039A-11731
18	426	10.5	521	4	US-09-489-039A-9549
19	421.5	10.3	584	2	US-08-928-692-13
20	421.5	10.3	584	3	US-09-339-972-13
21	384.5	9.4	534	2	US-09-031-392-4
22	384.5	9.4	534	3	US-09-299-549-4
23	384.5	9.4	534	3	US-09-610-417-4
24	365	9.0	493	2	US-09-031-392-10
25	365	9.0	493	3	US-09-299-549-10
26	365	9.0	493	3	US-09-610-417-10
27	358	8.8	488	2	US-08-928-692-10

28	358	8.8	488	3	US-09-339-972-10	Sequence 10, Appl
29	346	8.5	524	2	US-08-928-692-12	Sequence 12, Appl
30	346	8.5	524	3	US-09-339-972-12	Sequence 12, Appl
31	344	8.4	509	2	US-09-031-392-6	Sequence 6, Appl
32	344	8.4	509	3	US-09-299-549-6	Sequence 6, Appl
33	344	8.4	509	3	US-09-610-417-6	Sequence 6, Appl
34	341.5	8.4	504	4	US-09-679-686B-21	Sequence 21, Appl
35	340	8.3	511	4	US-09-679-686B-12	Sequence 12, Appl
36	331	8.1	492	2	US-08-355-844-3	Sequence 3, Appl
37	331	8.1	492	5	PCT-US95-16126-3	Sequence 3, Appl
38	330	8.1	488	2	US-08-928-692-11	Sequence 11, Appl
39	330	8.1	488	3	US-09-339-972-11	Sequence 11, Appl
40	328.5	8.1	863	4	US-09-894-927B-9	Sequence 9, Appl
41	322	7.9	552	4	US-09-248-796A-20713	Sequence 20713, A
42	315.5	7.7	514	4	US-09-679-686B-22	Sequence 22, Appl
43	312	7.7	502	4	US-09-679-686B-2	Sequence 2, Appl
44	310	7.6	494	2	US-09-031-392-5	Sequence 5, Appl
45	310	7.6	494	3	US-09-299-549-5	Sequence 5, Appl

ALIGNMENTS

RESULT 1
US-09-291-922-2
; Sequence 2, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 747
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (129)
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (133)..(134)
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (144)
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (178)
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (207)
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (218)
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (220)
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (236)
US-09-291-922-2

Query Match 86.3%; Score 3517; DB 3; Length 747;
Best local Similarity 91.6%; Pred. No. 0;
Matches 684; Conservative 27; Mismatches 34; Indels 2; Gaps 2;

[illegible]

RESULT 2

US-09-291-922-8

05 03 221 222
; Sequence 8, Application US/09291922

Patent No. 6383776

GENERAL INFORMATION:

APPLICANT: Allen, Steve

APPLICANT: Hitz, Bill

APPLICANT: Kinney, Tony

APPLICANT: Tingey, Scott

; AFFILIANT: LINNEY, SCOTT
; TITLE OF INVENTION: Plant Sugar Transport Proteins

FILE OF INVENTION: 1244
FILE REFERENCE: BB-1163

FILE REFERENCE: DD FORM 1003
: CURRENT APPLICATION NUMBER: US/09/291.922

CURRENT FILING DATE: 1999-04-14

CONCERN FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044

; EARLIER AFFILIATION NUMBER: 860/083;
: EARLIER FILING DATE: April 24, 1998

: NUMBER OF SEQ ID NOS: 30

; NUMBER OF SEQ ID NOS: 30
: SOFTWARE: Microsoft Office 97; SOFTWARE: 8
; SEQ ID NO 8

SEQ ID NO. 8
LENGTH: 737

; TYPE: PRT
; ORGANISM: Glycine max
US-09-291-922-8

Query Match 65.6%: Score 2674: DB 3: Length 737:

Query Match 03:0%, Score 2071, DE 3,
Best Local Similarity 69.9%:
Pred. No. 4.7e-256:

Best Local Similarity	93.3%	Freq: NO: 4:76 230;	
Matches	522:	Mismatches	115:
Conservative	92:	Indels	18:
Conservative	92:	Gaps	8:

QY	56	MSGAVLVAIVASIGNLLQGDWNATIAAAVLYIKKEFQLQNEPTVEGLIVMSMLIGATIVT	111
Db	1	MKGAVLVAIAASIGNFLQGDWNATIAGANGYIKKDALGT--TMERLVVGMSLIGATVIT	58
QY	116	TPSGPLSDSIGRRRMLILSSILYFFSGLIMLWSPNVYVLLIARFVDGFGIGLAVTLVPLY	175
Db	59	TCSGPIADWLGRRRPMIISSVLYFLGGLVMLWSPNVYVLCIARLLDGGIGLAVTLVPVY	118
QY	176	ISEIAPSEIRGLNLTLPPQFSGSGGMFLSYCMVFGMSLSPSPDWIRIMLGVAIPSLFFFG	235
Db	119	ISEIAPSEIRGLNLTLPPQFSGSGGMFLSYCMVFGMSLSPAPSWRLMLGLVLSIPSLLYFAL	178
QY	236	TIFYLPESPRWLVSKGRMAEAKKVLQKLRGKDDVSGELSLLLEGLEVGCGDTSIEEYIIGP	295
Db	179	TIFFLPESPRWLVSKGRMLEAKKVLQRLRGREDVSGEMALLVEGLIGCGDTSIEEYIIGP	238
QY	296	ATEAADDLVTDGKXEQITLYGPEEGQSWIARPSKGPIMLGSVLSLASRHGSMVNSQVPLM	355
Db	239	ADDVADGHEHATEKDKIRLYGSQAGLSWLSKPVGTQSSIG---LASHHGSIIINQSMPLM	294
QY	356	DPIVTLFGSVHENMPQAG--GSMRSTLFPNFGSMFVSVDQHAQNEQWDEENLHRDDEEYA	413
Db	295	DPLVTLFGSIHEKLPETGARGSMRSTLFPNFGSMFSAEPHAKIEQWDEESLQREDYDYM	354
QY	414	SDGAGGDYEDNLHSPILSRQATGAEGKDIVHHGHRGSAL-SMRRQTLLGEG-GDGVSSTD	471
Db	355	SDATRGDSDNLHSPILSRQTTSL-KDLPPTPPSHGSILGSMRRHSSLMQSGGEGQSGTG	413
QY	472	IGGGWQLAWKWEKEGENRKEGGFKRVYLHQEGVPSRRGSIVSLPGGGDVFEGSEFVH	531
Db	414	IGGGWQLAWKWTDK-GEDGKQGGFKRIYLHEEGVSAARRGSIVSIPGEG-----EFVQ	466
QY	532	AAALVSQSALFSKGLAEPMSDAAAMVHPSEVAAGSRWKDLFEPGVRRALLVGVGIQILQ	591
Db	467	AAALVSQPALYSKELIDGHPVGPAMVHPSETASKGSPWKALLEPGVKHALVVGVIQILQ	526
QY	592	QFAGINGVLYTTPQILEQAGVAVILSKFGLSSASASILISLTLMLPCIGFAMLLMDL	651
Db	527	QFSGINGVLYTTPQILEEAGVEVLLSDIGIGSESASFLISAFITFLMLPCIGVAMKLMV	586
QY	652	SRRRLLGTTPILIASLVILVNSLIDLGLTAHALLSTVSIVVYFCCFVMGFGPIPNIL	711
Db	587	SRRRQLLLTTPVLIVSLIILVIGSLVNFNGVAHAASITVCVVVYFCCFVMGYGPIPNIL	646
QY	712	CAEIFPTRVRGLCIACTFWIGDIIIVTYSLPVMLNAGLAGVFSIYAVVCLISFVFV	771
Db	647	CSEIFPTRVRGLCIAICALVFWIGDIIITYSLPVMLGSLGLGGVFAIYAVVCFISWIFV	706
QY	772	LKVPETKGMPLVITEFFAVGAKQAAA	798
Db	707	LKVPETKGMPLVISEFFSVGAKQAAAS	733

RESULT 3

US-09-291-922-29

03 03 201 022 23 : Sequence 29, Application US/092919222

Patent No. 6383776

TELEPHONE NO.: 3303173
: GENERAL INFORMATION:

APPLICANT: Allen, Steve

APPLICANT: Hitz, Bill

APPLICANT: KINNEV, TONY

APPLICANT: TINGEV, Scott

: TITLE OF INVENTION: Plant Sugar Transport Proteins

FILE OF INVENTION: F18
FILE REFERENCE: BB-1163

FILE REFERENCE: BB 1103
: CURRENT APPLICATION NUMBER: US/09/291.922

; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 29
; LENGTH: 729
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-291-922-29

Query Match 63.4%; Score 2584; DB 3; Length 729;
Best Local Similarity 68.8%; Pred. No. 4.1e-247;
Matches 512; Conservative 89; Mismatches 123; Indels 20; Gaps 8;

QY 56 MSGAVLVAIVASIGNLLQGWDNATIAAAVLYIKKEFQLONEPTVEGLIVMSLIGATIVT 115
Db 1 MSGAVLVAIAAAGVNLQGWDNATIAAGAVLYIKKEFNLESNPSVEGLIVAMSLIGATLIT 60
QY 116 TFSGPLSDSIGRRPMLILSSILYFFSGLIMLWSPNVYVLLARFVDFGFIGLAVTLVPLY 175
Db 61 TCSGGVADWLGRRPMLILSSILYFVGSVLMWSPNVYVLLLRLLDGFVGLVTLVPIY 120
QY 176 ISEIAPSEIRGLNLTLPQFSGGMLFLSYCMVFGMSLSPSPDWRIMLVLAIPSLFFGL 235
Db 121 ISETAPSEIRGLNLTLPQFTSGGMLFLSYCMVFGMSLSPSPWRMLGVLFIPSLVFFFL 180
QY 236 TIFYPESPRWLVSGRMAEAKVLQKRGKDDVSGELSLLLEGLVGGDTSEIYYIIGP 295
Db 181 TVFFLPESPRWLVSGRMLEAKRVLQRLRGREDVSGEMALLVEGLIGGETTIEYYIIGP 240
QY 296 ATEAADLVTGDKEQITLYGPEEGQSWIARPSKGPIMLSVLSASRHGS-MVNQSVPL 354
Db 241 ADEVTDHDIADVCKQIKLYGAEGLSWVARPVKG----GSTMSVLSRHGSTMSRRQGS 296
QY 355 MDPIVTLFGSVHENMPQAGGSMRSTLFPNFGSMFVTDQAHAKNEQWDEENLHRDDEEYAS 414
Db 297 IDPLVTLFGSVHEKMPDT-GSMRSALFPFHFGSMFVSGNQPRHEDWDEENLVGEDEDYPS 355
QY 415 DGAGGDYEDNLHSPLLSRQATGAEGKDIVHHGHRGSALSMMRQTLLGEGDGVSTDIGG 474
Db 356 D-HGDDSEDDLHSPILSRQTSME-KDMPHTAHGTLSTFRHGSQVQGAQEGAGSMGIGG 413
QY 475 GWQLAWKWEKEGEGKGGFKRVYLHQEGVPGSRRGSIVSLPGGGDVFESEFVHAAA 534
Db 414 GWQVANKWTEREDESQKE-----EGFPSSRRGSIVSLPGGDGTGE-ADFVQASA 462
QY 535 LVQSALFSKGLAEPRMSDAAMVHPSEVAAKGSRWKDLFEPGVRRALLVGVGIQILOQFA 594
Db 463 LVQPALYSKDLLKEHTIGPAMVHPSE-TTKGSIWHDLDHPGVKRALVVGVLQILOQFS 521
QY 595 GINGVLYYTPQILEQAGVAVILSKFGLSSASASILSSLTLLMLPCIGFAMLLMDLSGR 654
Db 522 GINGVLYYTPQILEQAGVGILLSNMGISSSSASLLISALTTFVMLPAIAVAMRLMDLSGR 581
QY 655 RFLLLGTIPILIASLVVSNLIDLGLTAHALLSTVSVIVYFCCFVMGFGPIPNILCAE 714
Db 582 RTLLLTIPILIASLLVLVSNLVHMNSIVHAVLSTVSVVLYFCFFVMGFGPAPNILCSE 641
QY 715 IFPTRVRGLCIAICAFTEFWIGDIIIVTYSLPVMLNAIGLAGVFSIYAVVCLISFVFLKV 774
Db 642 IFPTRVRGICIAICALTFWICDIIIVTYSLPVLLKSIGLAGVFGMYAIVCCISWVVFVFKV 701
QY 775 PETKGMPLLEVITEFFAVGAKQAAA 798
Db 702 PETKGMPLLEVITEFFSVGARQAEA 725

RESULT 4
US-09-291-922-10
; Sequence 10, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:

; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 10
; LENGTH: 486
; TYPE: PRT
; ORGANISM: Glycine max
US-09-291-922-10

Query Match 41.0%; Score 1670.5; DB 3; Length 486;
Best Local Similarity 68.1%; Pred. No. 1.1e-156;
Matches 333; Conservative 57; Mismatches 84; Indels 15; Gaps 8;

QY 308 DKEQITLYGPEEGQSWIARPSKGPIMLSVLSASRHGSMMNQSVPLMDPIVTLFGSVHE 367
Db 5 EKQIKLYGPEQGSWVARPVAGPNVSG---LVSRKGSMANPS-SLVDPLVTLFGSVHE 59
QY 368 NMPQAGGSMRSTLFPNFGSMFVTDQAHAKNEQWDEENLHRDDEEYASDGAGDYEDNLHS 427
Db 60 KLPETG---STLFPFHFGSMFVSGNQPRNEDWDEESLAREGDDYVSD--AGSDDNLQS 113
QY 428 PLLSRQATGAEGKDIVHHGHRGSALSMMRQTLL-GEAGDGVSTDIGGWQLAWKWEKE 486
Db 114 PLTSRQTSLD-KDIPPHAHSNLA-SMRQGSLLHGSSEPTGSTGIGGWQLAWKWSERE 171
QY 487 GENRKEGGFKRVYLHQEGVPGSRRGSIVSLPGGGDVFESEFVHAAALVSQSALFSKGL 546
Db 172 GPDGKEGGFKRIYHLHQDGGSGSRRGSVVSLLP-GGDLPTDSEVVQAAALVSQPALYNEDL 230
QY 547 AEPRMSDAAMVHPSEVAAKGSRWKDLFEPGVRRALLVGVGIQILOQFAGINGVLYYTPQI 606
Db 231 MRQPPVGPAMIHPSETIAKPSWSDLFEPGVKHALIVGVGMQILQQFSGINGVLYYTPQI 290
QY 607 LEQAGVAVILSKFGLSSASASILSSLTLLMLPCIGFAMLLMDLSGRRFLILGTIPILI 666
Db 291 LEQAGVGYLLSSLLGLGSTSSFLISAVTLLMLPCIAIAMRLMDISGRRTLLLSLTPVLI 350
QY 667 ASLVILVVSNLIDLGLTAHALLSTVSVIVYFCCFVMGFGPIPNILCAEIFFPTRVGLCIA 726
Db 351 AALLILVLSLVDLSTANASISTISIVIVYFCCFVMGFGPIPNILCAEIFFPTRVGLCIA 410
QY 727 ICATFTWIGDIIIVTYSLPVMLNAIGLAGVFSIYAVVCLISFVFLKVPETKGMPLVIT 786
Db 411 ICALTFWICDIIIVTYPVMLNSVGLAGVFGIYAVVCFIAWVFLKVPETKGMPLVIT 470
QY 787 EFFAVGAKQ 795
Db 471 EFFSVGAKQ 479

RESULT 5
US-09-291-922-14
; Sequence 14, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044

```

; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 14
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-09-291-922-14

```

	Query Match	34.9%;	Score 1424;	DB 3;	Length 345;
	Best Local Similarity	82.9%;	Pred. No. 1.9e-132;		
	Matches 281;	Conservative 24;	Mismatches 32;	Indels 2;	Gaps 2;
Qy	462	EGDGVSSTDIGGWQLAWKWKSEKEGENRKEGGFKRVYLHQEGVPGSRRGSIIVSLPGGG	521		
Dd	4	EGGEAVSSTGICGGWQLAWKWSERQGEDGKKEGGFKRIYVLHQEGVADSRRGSVVSLPGGG	63		
Qy	522	DVFE-GSEFVHAAALVSQSALFSKGLAEPRM-SDAAMVHPSEVAAGSRWKDLFEPGVRR	579		
Dd	64	DATQGSGGFIIHAAALVSHSALYSKDLMEERMAAGPAMIHFLEAAPKGSIWKDLFEPGVRR	123		
Qy	580	ALLVGVGIOILQQFAGINGVLYYTPOILEQAGVAVILSKPGLSSASASILSSLTTLML	639		
Dd	124	ALFVGVGIOMLQQFAGINGVLYYTPOILEQAGVAVLLSNLGLSSASASILSSLTTLML	183		
Qy	640	PCIGFAMLLMDLSGRFRLLLTGTIPILIASLVILVSNLIDLGTLAHALLSTVSVIVTFCC	699		
Dd	184	PSIGVAMRLMDISGRFRLLLTGTIPILIASLVLGVNVINLSTVPHAVLSTVSVIVTFCC	243		
Qy	700	FVMGFGPIPNILCAEIFFPTRVRGLCIAICFTFWICDIIVTYSLPVMLNAIGLAGVESIY	759		
Dd	244	FVMGFGPIPNILCAEIFFPTRVRGVCIACALTFWICDIIVTYSLPVMLNAIGLAGVFGIY	303		
Qy	760	AVVCLISFVFVLKVPETKGMPLLEVITEFFAVGAQAQAAA	798		
Dd	304	AVVCCIAFVFVLKVPETKGMPLLEVITEFFAVGAQAQAAA	342		

RESULT 6

```

US-09-291-922-16
; Sequence 16, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Tra
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/2
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 16
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-09-291-922-16

```

	Query Match	23.1%;	Score 943;	DB 3;	Length 228;	
	Best Local Similarity	80.6%;	Pred. No. 5.4e-85;			
	Matches 179; Conservative	25;	Mismatches 18;	Indels 0;	Gaps 0;	
QY	574	EPGVRALLVGVGIQLQQFAGINGVLVYTPQILEQAGVAVILSKFGLSSASASILISSL	633			
		:	:	:	:	:
		:	:	:	:	:
		:	:	:	:	:
		: :	:	:	:	:
		:	:	:	:	:
		:	:	:	:	:
Dd	1	EPGVKHALFVGIGLQILQQFAGINGVLVYTPQILEQAGVGVLLSNIGLSSSASILISAL	60			
		:	:	:	:	:
		:	:	:	:	:
		:	:	:	:	:
		: :	:	:	:	:
		:	:	:	:	:
		:	:	:	:	:
QY	634	TTLMLPCIGPAMLLMDLSGRREFLLGTTPIILIASLVILVVSNLIDLGTLAHALLLSTVSV	693			
		:	:	:	:	:
		:	:	:	:	:
		:	:	:	:	:
		: :	:	:	:	:
		:	:	:	:	:
		:	:	:	:	:
Dd	61	TTLMLPSIGIAMRLMDMSGRRFFLLSTIPVLIVVALAVLVNVLDVGTVMVAALSTISV	120			
		:	:	:	:	:
		:	:	:	:	:
		:	:	:	:	:
		: :	:	:	:	:
		:	:	:	:	:
		:	:	:	:	:

QY	694	I V Y F C F V M G F G P I P N I L C A E I P P T R V R G L C I A I C A L T F W I G D I I V T Y S L P V M L N A I G L A	753
D b	121	I V Y F C F V M G F G P I P N I L C A E I P P T S V R G I C I A I C A L T F W I G D I I V T Y T L P V M L N A I G L A	180
		:	
QY	754	G V F S I Y A V V C L I S F V F V L K V P E T K G M P L E V I T E F F A V G A K Q	795
		: : :	
D b	181	G V F G I A I V C V L A F V F V M K V P E T K G M P L E V I T E F F S V G A K Q	222

RESULT 7

```

US-09-291-922-6
; Sequence 6, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 6
; LENGTH: 131
; TYPE: PRT
; ORGANISM: Oryza sativa
US-09-291-922-6

```

Query Match	13.8%;	Score 562;	DB 3;	Length 131;
Best Local Similarity	80.5%;	Pred. No. 1.4e-47;		
Matches 103;	Conservative 17;	Mismatches 8;	Indels 0;	Gaps 0;
QY	669	LVLVVSNLIDLGTLAHALLSTVS	VIVYFCCFVMGFGPIPNILCAE	IFPTRVRLCIAIC 728
Db	1	VLTLLVNILDVGTMTVHASLSTVS	VILYFCFFVMGFGPIPNILCAE	IFPTRVRLCIAIC 60
QY	729	AFTFWIGDIIIVTYSLPVMLNAI	GLAGVFSIYAVVCLISFVFLKVP	ETKGMPLVITEF 788
Db	61	ALTFWIGDIIIVTYSLPVMLNAI	GLAGVFGIYAVVCILAFLFVFMKVP	ETKGMPLVITEF 120
QY	789	FAVGAKQA	796	
Db	121	FSVGAKQA	128	

RESIST 8

US-10-162-012-46
; Sequence 46, Application US/10162012
; Patent No. 6682597
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A.J.
; APPLICANT: Silos-Santiago, Inmaculada
; APPLICANT: Gu, Wei
; TITLE OF INVENTION: NOVEL HUMAN ION CHANNEL AND TRANSPORTER FAMILY MEMBERS
; FILE REFERENCE: 10448-190001
; CURRENT APPLICATION NUMBER: US/10/162,012
; CURRENT FILING DATE: 2002-06-04
; PRIOR APPLICATION NUMBER: US 60/209,845
; PRIOR FILING DATE: 2000-06-06
; PRIOR APPLICATION NUMBER: US 09/875,321
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: PCT/US01/18340
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/209,257
; PRIOR FILING DATE: 2000-06-05
; PRIOR APPLICATION NUMBER: US 09/875,423
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: PCT/US01/18398
; PRIOR FILING DATE: 2001-06-05

Db 66 SLYIKRDLQITDVQLEIMMGILSVYALIGSFL-----GARTSDWVGRRVTVVFAAAIFNNG 121

QY 142 GLIMLWSPNVVLLARFVDGFGIGLAVTLVPLYIYSEIAPSEIRGLLNTLPQFSGSGGMF 201

Db 122 SLLMGFAVNYAMLMVGRFVTGIGVGYAIMVAPVYTPEVSPASARGFLTSFTEVFINVGIL 181

QY 202 LSYCMVFGMSLSP-SPDWRIMLGVLAIPLSPFFGLTIFYLPESPRWLVSCKGRMAEAKKVL 260

Db 182 LGYVSNYAFARLPLHLWRVLMGIGAVPSA-LLALMVFGMPESPRWLVMKGRADARAVL 240

QY 261 QKLRGKDDVSGELSLLEGLEVGDDTSIEEYIIGPATEAADLVTGDKKEQITLYGPEEG 320

Db 241 ----- 240

QY 321 QSWIARPSKGPIMLGSVLASRHHGSMVNQSVPLMDPIVTLFGSVHENMPQAGSMRSTL 380

Db 241 ----- 240

QY 381 FPNFGSMFVTDQHAKNEQWDEENLHRDDEYASDGAGDYEDNLHSPLLSRQATGAEGK 440

Db 241 -----AKTSDTPEEAVERLDQIKAA----- 260

QY 441 DIVHHGRGSALSMMRQTLLGEGDGVSTDIGGWQLAWKWKSEKENGKRGKRVY 500

Db 261 ----- 260

QY 501 LHQEGVPSRRGSIIVSLPGGDDVFEGSEFVHAAALVSQALFSKGLAEPMSDAMVHPS 560

Db 261 ---AGIPRELDGDDVVMFK-----TKGQEKQV----- 285

QY 561 EVAAGSRWKDLF---EPGVRALLVGVGIQILQQFAGINGVLYYTPQILEQAGVAVILS 617

Db 286 -----WKELIFPTPAMRRILLAAALGIHFFQOATGSDSVLYSPRVFQSAGITGDNH 337

QY 618 KFGLSASASILISLTLMLPCIGFAMLLMDLSGRRFLLLTGTIPILLASLV-----IL 672

Db 338 LLG-----ATCAMGVMTLFI-----VATFQLDVRGRRPLLLTSTAGMLACLIGLGTGLT 388

QY 673 VVSNLIDLGTLAHALLSTVSVIVYFCCFVMGFGPIPNILCAEIFFPTRVRGLCIAICATF 732

Db 389 VVGRHPDAKVPWAIGLCIVSILAYVSFFSIGLPLTSVYTSVTFPLRVLRALGFALGTSCN 448

QY 733 WIGDIIIVTSLPVMNAIGLAGVFSIYAVVCLISFVFVFLKVPETKGMPLVITEFF--- 789

Db 449 RVTSAAVSMFSLSKAITIGGSFFLYAGIAAGIWIFFFTFIPETRGLPLEIGKLFGMT 508

QY 790 --AVGAKQAAK 799

Db 509 DTAVEAQDTATK 520

RESULT 12

US-09-291-922-30

; Sequence 30, Application US/09291922

; Patent No. 6383776

; GENERAL INFORMATION:

; APPLICANT: Allen, Steve

; APPLICANT: Hitz, Bill

; APPLICANT: Kinney, Tony

; APPLICANT: Tingey, Scott

; TITLE OF INVENTION: Plant Sugar Transport Proteins

; FILE REFERENCE: BB-1163

; CURRENT APPLICATION NUMBER: US/09/291,922

; CURRENT FILING DATE: 1999-04-14

; EARLIER APPLICATION NUMBER: 60/083,044

; EARLIER FILING DATE: April 24, 1998

; NUMBER OF SEQ ID NOS: 30

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 30

; LENGTH: 549

; TYPE: prt

; ORGANISM: Beta vulgaris

US-09-291-922-30

Query Match 12.0%; Score 488; DB 3; Length 549;

Best Local Similarity 21.3%; Pred. No. 4e-39;

Matches 164; Conservative 113; Mismatches 194; Indels 298; Gaps 19;

QY 41 SSQEPVTSDDILED-----KMSGAVLVAIVASIGNLLQGDNATIAAAVLYIKKEF 91

Db 10 SDPPPTTASKVIADFDPLKKPKRNKFAFACATLASMTSVLLGYDIGVMSGALYILKEDW 69

QY 92 QLQNEPTVEGLIVMSMLICATIVTTFSGPLSDSISGRPMILSSILYFFSGLIMLWSPNV 151

Db 70 HISD--TQIGVLVILNIYCLFGSFAAGRTSDWIGRRYTIIVLAGAIFVFGALLMGFATNY 127

QY 152 YVLLLARFVDGFGIGLAVTLVPLYIYSEIAPSEIRGLLNTLPQFSGSGMFLSYCMVFGMS 211

Db 128 AFLMVGRFVTGIGVGYALMIAPVYTAEVSPASSRGRFLTSFPEVFINAGILLGYISNLAFS 187

QY 212 LSPSP-DWRIMLGVLAIPLSPFFGLTIFYLPESPRWLVSCKGRMAEAKKVLQKRGKDDVS 270

Db 188 SLPTHLSWRFLGIGAIPSI-FLAIGVLAMPESPRWLVMQGRGLGDAKKVLNR----- 238

QY 271 GELSLLLEGEVGGDTSIEEYIIGPATEAADLVTGDKKEQITLYGPEEGQSWIARPSKG 330

Db 239 -----ISD-----SPEEAQ----- 247

QY 331 PIMLGSVLASRHHGSMVNQSVPLMDPIVTLFGSVHENMPQAGSMRSTLFPNFGSMFSV 390

Db 248 ----- 247

QY 391 TDQHAQNEQWDEENLHRDDEYASDGAGDYEDNLHSPLLSRQATGAEGKDIVHHGHRGS 450

Db 248 -----L 248

QY 451 ALSMRQTLLEGGDGVSTDIGGWQLAWKWKSEKENGKRGKRVYLHQEGVPSR 510

Db 249 RLSEIKQT-----AGIPA-- 261

QY 511 RGSIVSLPGGDDVFEGSEFVHAAALVSQALFSKGLAEPMSDAMVHPSVAAGSRWK 570

Db 262 -----ECDE-----DIYKVEKTKIKSGNAVWK 283

QY 571 DLF---EPGVRALLVGVGIQILQQFAGINGVLYYTPQILEQAGVAVILSKFGLSSASAS 627

Db 284 ELFFNPPTPAVRRAVIAGIGIHFFQOASGIDAVVLYSPRIFQSAGITNARKQL-----LAT 338

QY 628 ILISSLTLLMLPCIGFAMLLMDLSGRRFLLLTGTIPILLASLVILVVS-NLIDLG-----T 682

Db 339 VAVGVVKTFLFI---VATFQLDKYGRRPLLLTSTAGMLACLIGLGTGLT 394

QY 683 LAHALLSTVSVIVYFCCFVMGFGPIPNILCAEIFFPTRVR--GLCIAICATFVIGDIIVT 740

Db 395 WAIALCITM-VCAVWASFSIGLGPITWVYSSEVFPLRLRAQGTSMGV-AVNRVVGVISI 452

QY 741 YSLPVMNAIGLAGVFSIYAVVCLISFVFVFLKVPETKGMPLVITEFF 789

Db 453 FFLLP-LSHKITTGGAFFLFGGIAIAIAWFFFTFLPETRGRLENMHLEF 500

RESULT 13

US-09-291-922-22

; Sequence 22, Application US/09291922

; Patent No. 6383776

; GENERAL INFORMATION:

; APPLICANT: Allen, Steve

; APPLICANT: Hitz, Bill

; APPLICANT: Kinney, Tony

; APPLICANT: Tingey, Scott

; TITLE OF INVENTION: Plant Sugar Transport Proteins

; FILE REFERENCE: BB-1163

; CURRENT APPLICATION NUMBER: US/09/291,922

; CURRENT FILING DATE: 1999-04-14

; EARLIER APPLICATION NUMBER: 60/083,044

; EARLIER FILING DATE: April 24, 1998

; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 22
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (102)
US-09-291-922-22

Query Match 11.9%; Score 485; DB 3; Length 510;
Best Local Similarity 22.8%; Pred. No. 7e-39;
Matches 173; Conservative 98; Mismatches 187; Indels 300; Gaps 19;

QY 59 AVLVAIVASIGNLLQWDNATIAAAVLYIKKEFQLQNEPTVE---GLIVSMSLIGATIVT 115
Db 21 AFACAILASMTSILLGYDIGVMSGASLYIKKDFNI-SDGKVEVLMGILNLYSLIGSFA-- 77
QY 116 TFSGPLSDSISGRPRMLILSSILYFFSGGLIMLWSPNVYVLLARFVDGFGIGLAVTLVPLY 175
Db 78 --AGRTSDWIGRRYTIFFAAVIFAGXFLMGFAVNYAMLMFGRFVAGIGVGYALMIAPVY 135
QY 176 ISEIAPSEIRGLNLTLPQFSGSGMFLSYCMVFGMSLSP-SPDWRLMGLVLAIPSLFFFG 234
Db 136 TAEVSPASARGFLTSPPEVFINFGLLGYVSNYAFSLPLNLGWRIMLIGIGAAPSV-LLA 194
QY 235 LTIFYLPESPRWLYSKGRMAEAKKVLQKRGKDDVSGELSLLLEGLLEVGGDTISIEEYIIG 294
Db 195 LMVLGMPESPRWLVMKGLADAKVLEKT-----SDT----- 226
QY 295 PATEAADDLVTGDKEQITLYGPEEGQSWIARPSKGPIMLGSVLSLASRHGSMVNQSVPL 354
Db 227 -AEEAAERL----- 234
QY 355 MDPIVTLFGSVHNNMPQAGGSMRSTLFPNFGSMFVTDQHAKEQWDEENLHRDDEEYAS 414
Db 235 ----- 234
QY 415 DGAGGDYEDNLHSPLLSRQATGAEGKDIVHHGHRGSALSMMRQTLLGEGGDVSVSTDIGG 474
Db 235 -----ADIK 239
QY 475 GWQLAWKWKSEKENGKRGKGGFKRVYLHQEGVPGSRRGSIIVSLP--GGGDVFESEFVHA 532
Db 240 A-----AGIPEELDGDVVTVPKRGSGN----- 261
QY 533 AALVSQSALFSKGLAEPRMSDAAAMVHPSEVAAGSRWKOLF---EPGVRRALLVGVGIQI 589
Db 262 -----EKR-----WKELILSPTPAMRRILLSGIGIHF 289
QY 590 LQQFAGINGVLYYTPQILEQAGVAVILSKFGLSSASASILISSLTLLMLPCIGFAMLLM 649
Db 290 FQHALGIHSVVFYSPVFKSPGLT--NDKHFLGTTWPFVGTKRLLFILL-----ATFFI 340
QY 650 DLSGRRFLLLTIPILIASLV-----ILVSNLIDLGLTAHALLSTVSVIVYFCCFVMGF 704
Db 341 DGVRRPRLLLGSGIILSLIGLAGLTVVGQHPDAKI PWAIGLSIASTLAYVAFFSIGL 400
QY 705 GPIPNILCAIFPTRVR--GLCIAICFTFWIGDIIIVTSLPVMNLNAIGLAGVFSIYAWV 762
Db 401 GPITWVYSSBIFPLQVRALGCSLGVAAARVTSGVISMTFL--SLSKAITIGGSFFLYSGI 458
QY 763 CLISFVFLKVPETKGMPLVITEFFFAVGAQAQAAKA 800
Db 459 AALAWVFFYTYLPETRGRTLEEMSKLF--GDTAAASES 494

RESULT 14

US-09-291-922-26
; Sequence 26, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:

; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 26
; LENGTH: 539
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-09-291-922-26

Query Match 11.4%; Score 463.5; DB 3; Length 539;
Best Local Similarity 21.7%; Pred. No. 1.1e-36;
Matches 166; Conservative 108; Mismatches 200; Indels 291; Gaps 17;

QY 52 LEDKMSG-----AVLVAIVASIGNLLQWDNATIAAAVLYIKKEFQLQNEPTVE---GLIV 104
Db 35 VEPKKGNVREFACAILASMTSILLGYDIGVMSGASLYIQDLKI-NDTQLEVLMGILN 93
QY 105 SMSLIGATIVTFSGPLSDSISGRPRMLILSSILYFFSGGLIMLWSPNVYVLLARFVDGFG 164
Db 94 VYSLIGSFA---AGRTSDWIGRRYTIFFAAVIFAGALIMGFSVNYAMLMFGRFVAGIG 149
QY 165 IGLAVTLVPLYISEIAPSEIRGLNLTLPQFSGSGMFLSYCMVFGMS-LSPSPDWRLMGL 223
Db 150 VGYALMIAPVNTGEVSPASARGVLTSPPEVFINFGLLGYVSNFAPAFARLSRLGWRIMLG 209
QY 224 VLAIPSLFFFGLTIFYLPESPRWLVSGRMAEAKKVLQKRGKDDVSGELSLLLEGLEVG 283
Db 210 IGAVPSV-LLAFMVLGMPESPRWLVMKGLADAKVLAKT----- 248
QY 284 GDTISIEEYIIGPATEAADDLVTGDKEQITLYGPEEGQSWIARPSKGPIMLGSVLSLASR 343
Db 249 SDT-----PEEAERIA----- 260
QY 344 HGSMVNQSVPLMDPIVTLFGSVHNNMPQAGGSMRSTLFPNFGSMFVTDQHAKEQWDEE 403
Db 261 ----- 260
QY 404 NLHRDDEEYASDGAGGDYEDNLHSPLLSRQATGAEGKDIVHHGHRGSALSMMRQTLLGEG 463
Db 261 -----DIKTA----- 266
QY 464 GDGVSTDIGGQWLAWKWKSEKENGKRGKGGFKRVYLHQEGVPGSRRGSIIVSLPGGGDV 523
Db 267 -----GIPGLDGDVVVPVK----- 281
QY 524 FEGSEFVHAAALVSQSALFSKGLAEPRMSDAAAMVHPSEVAAGSRWKDLFEPGVRRALLV 583
Db 282 -----NKGSEKRVLKDILSPITIA-----MRHILIA 309
QY 584 GVGIIQLQFAGINGVLYYTPQILEQAGVAVILSKFGLSS-ASASILSSLTLLMLPCI 642
Db 310 GIGIHFFQSSGIDAVVLYSPLVFKSAGIT-----GDSRLRGTTVAGATNTVFIL--- 360
QY 643 GFAMLLMDLSGRRFLLLTIPILIASLV-----ILVSNLIDLGLTAHALLSTVSVIVYF 697
Db 361 -VATFLDRIRRRPLVLTSTGMLVSLVGLATGLTVISRHPDEKITWAILVCIFCIMAIV 419
QY 698 CCFVMGEGPIPNILCAEIPTRVRGLCIAICFTFWIGDIIIVTSLPVMNLNAIGLAGVFS 757
Db 420 AFFSIGLGPITWVYSSBIFPLHVRALGCSLGVAAARVTSGVISMTFISLSKAMTIGGAFF 479
QY 758 IYAVVCLISFVFLKVPETKGMPLVITEFFFAVGA--KQAAAKA 800
Db 480 LFAGIASFAWVFFFAVLPETRGRTLEDMSSLSFGNTATHKQGAEEA 524

QY 746 MLNAIGLAGVFSIYAVVCLISFVFLKVPETKGMPLVI 785
Db 420 MLNSIGSANTFWVYGLNVLFILLTLWLIPETKNVSLHI 459

Search completed: October 13, 2004, 11:40:10
Job time : 30.8337 secs

RESULT 15
US-09-489-039A-11933
; Sequence 11933, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 11933
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-11933

Query Match 11.0%; Score 447; DB 4; Length 476;
Best Local Similarity 20.8%; Pred. No. 3.7e-35;
Matches 158; Conservative 98; Mismatches 186; Indels 318; Gaps 18;
QY 42 SQEPVTS-DDILED-KMSG-----AVLVAIVASIGNLLQGDNATIAAAVLYIKKEFQL 93
Db 2 STESITQLEGVMPDNKQGRSNKTMFTFFVFLAALAGLLFGLDIGIAGALPFIANEFOI 61
QY 94 QNEPTVEGLIVMSLIGATIVTTFSGPLSDSIGRRPMLILSSILYFFSGLIMLWSPNVYV 153
Db 62 SAH--TQEWVSSMMFGAAGVAVGSGWLSFKLGRKKSMLGAILFVAGSLFSAAPNVEI 119
QY 154 LLLARFVDGFGIGLAVTLVPLYISEIAPSEIRGLNLTLPQFSGGGMFLSYCMVFGMSLS 213
Db 120 LLVSRVLLGLAVGASVAPLYLSEIAPEKIRGSMISMYQLMITIGILGAY--LSDTAFS 177
QY 214 PSPDWRIMGLVLAIPSLFFFGTLTFYLPESRWLVSKGRMAEAKKVLQKLRGKDDVSGEL 273
Db 178 YSGAWRWMLGVIIIPAVLLL-IGVIFLPDSRWFPAKRRFVDAERVLLRLR----- 227
QY 274 SLLLEGLVGGDTSIEEYIIGPATEAADLVTGDKEQITLYGPEGQSWIARPSKGPIM 333
Db 228 -----DTSAE----- 232
QY 334 LGSVLSLASRHGSMVNSQSVPLMDPIVTLFGSVHNNPQAGGSMRSTLFPNFGSMFSVTDQ 393
Db 233 ----- 232
QY 394 HAKNEQWDEENLHRDDEYASDGAGGYEDNLHSPLSRQATGAEGKDIVHHGRGSALS 453
Db 233 -AKREL-----DEIRESLKVKQS----- 249
QY 454 MRROTLLGEGDGVSTDIGGGWQLAWKWEKEGNGRKEGGFKRVYLHQEGVPGSRRGS 513
Db 250 -----GWSL-----FK----- 255
QY 514 IVSLPGGDDVFEGSEFVHAAALVQSALFSKGLAEPRMSDAAAMVHPSEVAAKGRWKDLF 573
Db 256 -----DNSNF----- 260
QY 574 EPGVRRALLVGVGIQILQQFAGINGVLYYTPQILEQAGVAVILSKFGLSSASASILISSL 633
Db 261 ----RRAVFLGILLQVMQQTGMNVIMYAPKIFELAGYANTTEQMW-----GTIVVGLT 311
QY 634 TTLMLPCIGFAMLLMDLSGRFLLLTGTPILIASLVILVNSNLIDLGTLAH----- 685
Db 312 NVLATFTIAG---LVDRWGRK-----PTLILGFIVMAAGMGV-LGTMMHIGHSSTA 359
QY 686 ALLSTSVIVYFCCFVMGFGPIPNILCAEIPFTRVRGLCIACAFTWIGDIIIVTYSLPV 745
Db 360 QYIAVLMMLMFIVGFAMSAGPLLIWLCSEIQPLKGRDFGTCSTATNWIANMIVGATFLT 419

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2004, 11:38:38 ; Search time 302.735 Seconds
(without alignments)
852.715 Million cell updates/sec

Title: US-10-051-909-32

Perfect score: 4075

Sequence: 1 IRSGSWLAVQTPTDRLRR.....PLEVITEFFAVGAKQAAKA 800

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1356558 seqs, 322682953 residues

Total number of hits satisfying chosen parameters: 1356558

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09D_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10F_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US10G_PUBCOMB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US10H_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	4075	100.0	800	US-10-051-909-32	Sequence 32, Appl
2	3774	92.6	745	US-10-310-154-718	Sequence 718, App
3	3517	86.3	747	US-10-051-902-2	Sequence 2, Appli
4	3517	86.3	747	US-10-051-909-2	Sequence 2, Appli
5	2869	70.4	564	US-10-425-114-56898	Sequence 56898, A
6	2834	69.5	740	US-10-051-909-37	Sequence 37, Appl
7	2834	69.5	740	US-10-437-963-195955	Sequence 195955,
8	2674	65.6	737	US-10-051-902-8	Sequence 8, Appli
9	2674	65.6	737	US-10-051-909-8	Sequence 8, Appli
10	2584	63.4	729	US-10-051-902-29	Sequence 29, Appl
11	2584	63.4	729	US-10-051-909-29	Sequence 29, Appl
12	2581.5	63.3	742	US-10-424-599-153166	Sequence 153166,
13	2543	62.4	728	US-10-425-114-47258	Sequence 47258, A
14	2496.5	61.3	742	US-10-424-599-251277	Sequence 251277,

15	2442.5	59.9	661	15	US-10-425-114-37283	Sequence 37283, A
16	2305	56.6	735	16	US-10-437-963-149907	Sequence 149907,
17	1963	48.2	652	16	US-10-437-963-195442	Sequence 195442,
18	1763	43.3	714	16	US-10-437-963-173392	Sequence 173392,
19	1670.5	41.0	486	13	US-10-051-902-10	Sequence 10, Appl
20	1670.5	41.0	486	13	US-10-051-909-10	Sequence 10, Appl
21	1557	38.2	648	15	US-10-424-599-220809	Sequence 220809,
22	1424	34.9	345	13	US-10-051-902-14	Sequence 14, Appl
23	1424	34.9	345	13	US-10-051-909-14	Sequence 14, Appl
24	1293	31.7	272	16	US-10-767-701-44159	Sequence 44159, A
25	1252.5	30.7	338	15	US-10-424-599-184842	Sequence 184842,
26	944	23.2	248	15	US-10-424-599-243978	Sequence 243978,
27	943	23.1	228	13	US-10-051-902-16	Sequence 16, Appl
28	943	23.1	228	13	US-10-051-909-16	Sequence 16, Appl
29	813	20.0	161	15	US-10-425-114-69883	Sequence 69883, A
30	793	19.5	158	15	US-10-425-114-47887	Sequence 47887, A
31	663	16.3	145	16	US-10-767-701-50386	Sequence 50386, A
32	626	15.4	218	16	US-10-437-963-162571	Sequence 162571,
33	620	15.2	126	16	US-10-767-701-44178	Sequence 44178, A
34	587.5	14.4	457	14	US-10-369-493-23324	Sequence 23324, A
35	562	13.8	131	13	US-10-051-902-6	Sequence 6, Appli
36	562	13.8	131	13	US-10-051-909-6	Sequence 6, Appli
37	554.5	13.6	487	9	US-09-795-693-27	Sequence 27, Appl
38	554.5	13.6	487	13	US-10-095-139-14	Sequence 14, Appl
39	554.5	13.6	487	14	US-10-156-239-27	Sequence 27, Appl
40	554.5	13.6	487	14	US-10-199-485-27	Sequence 27, Appl
41	554.5	13.6	487	14	US-10-391-399-45	Sequence 45, Appl
42	554	13.6	143	15	US-10-425-114-55090	Sequence 55090, A
43	549	13.5	501	15	US-10-424-599-239553	Sequence 239553,
44	548	13.4	488	10	US-09-794-822-11	Sequence 11, Appl
45	548	13.4	488	13	US-10-094-059-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1

US-10-051-909-32
; Sequence 32, Application US/10051909
; Publication No. US20020199217A1

GENERAL INFORMATION:

APPLICANT: Allen, Steve
APPLICANT: Helentjaris, Tim
APPLICANT: Hitz, Bill
APPLICANT: Kinney, Tony
APPLICANT: Tingey, Scott
TITLE OF INVENTION: Plant Sugar Transport Proteins
FILE REFERENCE: BB1163 US CIP
CURRENT APPLICATION NUMBER: US/10/051,909
CURRENT FILING DATE: 2002-01-17
PRIOR APPLICATION NUMBER: 60/083,044
PRIOR FILING DATE: April 24, 1998
NUMBER OF SEQ ID NOS: 38
SOFTWARE: Microsoft Office 97
SEQ ID NO 32
LENGTH: 800
TYPE: PRT
ORGANISM: Zea mays

US-10-051-909-32

Query Match 100.0%; Score 4075; DB 13; Length 800;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 800; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	IRSGSWLAVQTPTDRLRRRLPSVVLALPGPLPPASCSSQEPVTSDDILEDKMSGAV	60
Db	1	IRSGSWLAVQTPTDRLRRRLPSVVLALPGPLPPASCSSQEPVTSDDILEDKMSGAV	60
QY	61	LVAIVASIGNLLQGWDNATIAAAVLYIKKEFQLONEPTVEGLIVMSLIGATIVTTFSGP	120
Db	61	LVAIVASIGNLLQGWDNATIAAAVLYIKKEFQLONEPTVEGLIVMSLIGATIVTTFSGP	120
QY	121	LSDSIGRRRPMILSSILYFFSGLIMLWSPNVVLLARFVDGFGIGLAVTLVPLYISEIA	180

Db 121 LSDSIGRRPMLILSSILYFFSGLIMLWSPNVVLLARFVDGFGIGLAVTLVPLYISEIA 180
QY 181 PSEIRGLLNTLTPQFSGSGMFLSYCMVFGMSLSPSPDWIRMLGVLAIPSLFFGLTIFYL 240
Db 181 PSEIRGLLNTLTPQFSGSGMFLSYCMVFGMSLSPSPDWIRMLGVLAIPSLFFGLTIFYL 240
QY 241 PESPRWLVSCKGRMAEAKKVLQKLRGKDDVSGELSLLEGLVGGDTSEIEYIIGPATEAA 300
Db 241 PESPRWLVSCKGRMAEAKKVLQKLRGKDDVSGELSLLEGLVGGDTSEIEYIIGPATEAA 300
QY 301 DDLVTDGDKQITLYGPEEGQSWIARPSKGPIMLGSVLSASRHGSMVNSQSVPLMDPIVT 360
Db 301 DDLVTDGDKQITLYGPEEGQSWIARPSKGPIMLGSVLSASRHGSMVNSQSVPLMDPIVT 360
QY 361 LFGSVHENMPQAGGSMRSTLFPNFGSMFVTDQHAQNEQWDEENLHRDDEEYASDGAGGD 420
Db 361 LFGSVHENMPQAGGSMRSTLFPNFGSMFVTDQHAQNEQWDEENLHRDDEEYASDGAGGD 420
QY 421 YEDNLHSPLLSRQATGAEGKDIVHHGHRGSAISMRRQTLLEGEGDVSSTDIGGGWQLAW 480
Db 421 YEDNLHSPLLSRQATGAEGKDIVHHGHRGSAISMRRQTLLEGEGDVSSTDIGGGWQLAW 480
QY 481 KWSEKEGENRKEGFKRVYLHQEGVFGSGRRGSIIVSLPGGDDVFESEFVHAAALVSQSA 540
Db 481 KWSEKEGENRKEGFKRVYLHQEGVFGSGRRGSIIVSLPGGDDVFESEFVHAAALVSQSA 540
QY 541 LFSKGLAEPRMSDAAMVHPSEVAAGKSRWKDLFEPGVRRALLVGVIGIQLQOFAGINGVL 600
Db 541 LFSKGLAEPRMSDAAMVHPSEVAAGKSRWKDLFEPGVRRALLVGVIGIQLQOFAGINGVL 600
QY 601 YYPQILEQAGVAVILSKFGLSSASASILSSLTLLMLPCIGFAMLLMDLSGRRFLLLG 660
Db 601 YYPQILEQAGVAVILSKFGLSSASASILSSLTLLMLPCIGFAMLLMDLSGRRFLLLG 660
QY 661 TIPILIASLVILVNSLIDGLTLAHLSTVSIVVYFCCFVMGFGPIPNILCAEIFFTRV 720
Db 661 TIPILIASLVILVNSLIDGLTLAHLSTVSIVVYFCCFVMGFGPIPNILCAEIFFTRV 720
QY 721 RGLCIACAFTWIGDIIIVTYSPLVNLNAGLGVFSIYAVVCLISFVFLKVPETKGM 780
Db 721 RGLCIACAFTWIGDIIIVTYSPLVNLNAGLGVFSIYAVVCLISFVFLKVPETKGM 780
QY 781 PLEVITEFFFAVCAKQAAKA 800
Db 781 PLEVITEFFFAVCAKQAAKA 800

RESULT 2
US-10-310-154-718
; Sequence 118 Application US/10310154
; Publication No. US20030233670A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; APPLICANT: Chomet, Paul S.
; APPLICANT: Adams, Thomas H
; APPLICANT: Ruff, Thomas G.
; APPLICANT: Agarwal, Ameeta K.
; APPLICANT: Ahrens, Jeffrey E.
; APPLICANT: Ball, James A.
; APPLICANT: Banu, G.
; APPLICANT: Bell, Erin
; APPLICANT: Boddupalli, Raghava
; APPLICANT: Deikman, Jill
; APPLICANT: Deng, Molian
; APPLICANT: Dong, Jinzhuo
; APPLICANT: Duff, Stephen M.
; APPLICANT: Galligan, Meghan M.
; APPLICANT: Hinchey, Brenda S.
; APPLICANT: Huang, Shihshieh
; APPLICANT: Johnson, G. Richard
; APPLICANT: Jung, Vincent
; APPLICANT: Kretzmer, Keith A

; APPLICANT: Laccetti, Lucille B.
; APPLICANT: Lai, Chao-Qiang
; APPLICANT: Lee, Gary
; APPLICANT: Lin, Jie-Yi
; APPLICANT: Liu, Jingdong
; APPLICANT: Lu, Bin
; APPLICANT: Luethy, Michael M.
; APPLICANT: Lund, Adrian
; APPLICANT: Madson, Linda L.
; APPLICANT: Malloy, Kathleen A.
; APPLICANT: McKiel, Christine L.
; APPLICANT: Miller, Philip W.
; APPLICANT: Padmavathi, Manchikanti
; APPLICANT: Parnell, Laurence D.
; APPLICANT: Start, William G.
; APPLICANT: Tennesen, Dan
; APPLICANT: Vidya, K.R.
; APPLICANT: Wang, Haiyun
; APPLICANT: Xu, Zhanquo
; APPLICANT: Yang, Chunzhi
; APPLICANT: Zeng, Xiaoping
; APPLICANT: Zhang, Qiang
; APPLICANT: Zhao, Yajuan
; APPLICANT: Zhou, Li
; TITLE OF INVENTION: Gene Sequences and Uses Thereof in Plants
; FILE REFERENCE: 38-15(52796)B
; CURRENT APPLICATION NUMBER: US/10/310,154
; CURRENT FILING DATE: 2002-12-04
; PRIOR APPLICATION NUMBER: 60/337,358
; PRIOR FILING DATE: 2001-12-04
; NUMBER OF SEQ ID NOS: 736
; SEQ ID NO 718
; LENGTH: 745
; TYPE: PRT
; ORGANISM: Zea mays
US-10-310-154-718

Query Match 92.6%; Score 3774; DB 14; Length 745;
Best Local Similarity 99.6%; Pred. No. 1e-315;
Matches 742; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 56 MSGAVLVAIVASIGNLLQGWDNATIAAAVLYIKKEFQIQNEPTVEGLIVMSLIGATIVT 115
Db 1 MSGAVLVAIVASIGNLLQGWDNATIAAAVLYIKKEFQIQNEPTVEGLIVMSLIGATIVT 60
QY 116 TFSGPLSDSIGRRPMLILSSILYFFSGLIMLWSPNVVLLARFVDGFGIGLAVTLVPLY 175
Db 61 TFSGPLSDSIGRRPMLILSSILYFFSGLIMLWSPNVVLLARFVDGFGIGLAVTLVPLY 120
QY 176 ISEIAPSEIRGLLNTLTPQFSGSGMFLSYCMVFGMSLSPSPDWIRMLGVLAIPSLFFGL 235
Db 121 ISEIAPSEIRGLLNTLTPQFSGSGMFLSYCMVFGMSLSPSPDWIRMLGVLAIPSLFFGL 180
QY 236 TIFYLPESPRWLVSCKGRMAEAKKVLQKLRGKDDVSGELSLLEGLVGGDTSEIEYIIGP 295
Db 181 TIFYLPESPRWLVSCKGRMAEAKKVLQKLRGKDDVSGELSLLEGLVGGDTSEIEYIIGP 240
QY 296 ATEAADLVTDGDKQITLYGPEEGQSWIARPSKGPIMLGSVLSASRHGSMVNSQSVPLM 355
Db 241 ATEAADLVTDGDKQITLYGPEEGQSWIARPSKGPIMLGSVLSASRHGSMVNSQSVPLM 300
QY 356 DPVITLFGSVHENMPQAGGSMRSTLFPNFGSMFVTDQHAQNEQWDEENLHRDDEEYASD 415
Db 301 DPVITLFGSVHENMPQAGGSMRSTLFPNFGSMFVTDQHAQNEQWDEENLHRDDEEYASD 360
QY 416 GAGDYEDNLHSPLLSRQATGAEGKDIVHHGHRGSAISMRRQTLLEGEGDVSSTDIGGG 475
Db 361 GAGDYEDNLHSPLLSRQATGAEGKDIVHHGHRGSAISMRRQTLLEGEGDVSSTDIGGG 420
QY 476 WQLAWKWEKEGENRKEGFKRVYLHQEGVFGSGRRGSIIVSLPGGDDVFESEFVHAAAL 535
Db 421 WQLAWKWEKEGENRKEGFKRVYLHQEGVFGSGRRGSIIVSLPGGDDVFESEFVHAAAL 480

QY 536 VSQSALFSKGLAEPRMSDAAMVHPSEVAAGSRWKDLFEPGVRRLVGVGIQILQQFAG 595
DB 481 VSQSALFSKGLAEPRMSDAAMVHPSEVAAGSRWKDLFEPGVRRLVGVGIQILQQFAG 540
QY 596 INGVLYYTPQILEQAGVAVILSKFGLSSASASILISSLTTLMLPCIGFAMLLMDLSGRR 655
DB 541 INGVLYYTPQILEQAGVAVILSKFGLSSASASILISSLTTLMLPCIGFAMLLMDLSGRR 600
QY 656 FLLLGITIPILIASLVILVVSNNLIDGLTLAHLASTVSVIVYFCCFVMGFGPIPNILCAEI 715
DB 601 FLLLGITIPILIASLVILVVSNNLIDGLTLAHLASTVSVIVYFCCFVMGFGPIPNILCAEI 660
QY 716 FPTRVRGLCIAICAFTEFWIGDIIIVTYSPLVMLNAIGLAGVFSIYAVVCLISFVFVFLKVP 775
DB 661 FPTRVRGLCIAICAFTEFWIGDIIIVTYSPLVMLNAIGLAGVFSIYAVVCLISFVFVFLKVP 720
QY 776 ETKGMPLEVITEFFAVGAKQAAKA 800
DB 721 ETKGMPLEVITEFFAVGAKQAAKA 745

RESULT 3

US-10-051-902-2
; Sequence 2, Application US/10051902
; Publication No. US20020178468A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/10/051,902
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-14
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2

; LENGTH: 747
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: UNSURE
; NAME/KEY: (129)
; NAME/KEY: UNSURE
; LOCATION: (133)..(134)
; NAME/KEY: UNSURE
; LOCATION: (144)
; NAME/KEY: UNSURE
; LOCATION: (178)
; NAME/KEY: UNSURE
; LOCATION: (207)
; NAME/KEY: UNSURE
; LOCATION: (218)
; NAME/KEY: UNSURE
; LOCATION: (220)
; NAME/KEY: UNSURE
; LOCATION: (236)
US-10-051-902-2

Query Match 86.3%; Score 3517; DB 13; Length 747;
Best Local Similarity 91.6%; Pred. No. 1.4e-293;
Matches 684; Conservative 27; Mismatches 34; Indels 2; Gaps 2;

QY 56 MSGAVLVAIVASIGNLLQGWDNATIAAAVLYIKKEFQIQNEPTVEGLIVSMISIGATIVT 115
DB 1 MGGAVMVAIAASIGNLLQGWDNATIAAGVLYIKKEFNQSEPLIEGLIVAMFLIGATVIT 60
QY 116 TSGPLSDSIGRRPMLILSSILYFFSGLIMLWSPNVYVLLARFVDGFGIGLAVTLVPLY 175
DB 61 TSPGRADCVGRRPMLVASAVLYFVSGLVMLWAPIVYILLARLIDGFGIGLAVTLVPLY 120

QY 176 ISEIAPSEIR-GLLNTLPQFSG-SGGMFLSYCMVFGMSLSPSPDWRIMLVLAIPSLFFF 233
DB 121 ISETAPHRXSWGXXNTLPQFIGVXGGMFLSYCMVFGMSLMPKPDWRMLMLGVLSLPSLXYF 180
QY 234 GLTIFYLPESPRWLVSGRMAEAKVQLKRGKDDVSGELSLLLEGLVGGDTSEIYYII 293
DB 181 GLTVFYLPESPRWLVSGRMAEAKRVXQRLRGREDVSXEXALLVEGLGVGKDTRIKEYII 240
QY 294 GPATEAADDLVTDGKEQITLYGPEEGQSWIARPSKGPIMLGSVLSASRHGSMVNSQVSP 353
DB 241 GPATEAADDLVTDGKEQITLYGPEEGQSWIARPSKGPIMLGSVLSASRHGSMVNSQVSP 300
QY 354 LMDPIVTLFGSVHENMPOAGGSMRSTLFPNFGSMFVTDQHAQNEQWDEENLHRDDEEYA 413
DB 301 LMDPIVTLFGSVHENMPOAGGSMRSTLFPNFGSMFVTDQHAQNEQWDEENLHRDDEEYA 360
QY 414 SDGAGGDYEDNLHSPLLSROATGAEGKDIVHHGHRGSALSMMRRTLLGEGGDGVSSTDIG 473
DB 361 SDGAGGDYEDNLHSPLLSROATGAEGKDIVHHGHRGSALSMMRRTLLGEGGDGVSSTDIG 420
QY 474 GGWQLAWKWKSEKEGNGRKEGGFKRVYLHQEGVPGSRGSIIVSLPGGGDVFESEFVHAA 533
DB 421 GGWQLAWKWKSEKEGNGRKEGGFKRVYLHQEGVPGSRGSIIVSLPGGGDVFESEFVHAA 480
QY 534 ALVSQSALFSKGLAEPRMSDAAMVHPSEVAAGSRWKDLFEPGVRRLVGVGIQILQQF 593
DB 481 ALVSQSALFSKGLAEPRMSDAAMVHPSEVAAGSRWKDLFEPGVRRLVGVGIQILQQF 540
QY 594 AGINGVLYYTPQILEQAGVAVILSKFGLSSASASILISSLTTLMLPCIGFAMLLMDLSG 653
DB 541 AGINGVLYYTPQILEQAGVAVILSKFGLSSASASILISSLTTLMLPCIGFAMLLMDLSG 600
QY 654 RRFLLLGTIPILIASLVILVVSNNLIDGLTLAHLASTVSVIVYFCCFVMGFGPIPNILCA 713
DB 601 RRFLLLGTIPILIASLVILVVSNNLIDGLTLAHLASTVSVIVYFCCFVMGFGPIPNILCA 660
QY 714 EIFPTRVRGLCIAICAFTEFWIGDIIIVTYSPLVMLNAIGLAGVFSIYAVVCLISFVFVFLK 773
DB 661 EIFPTRVRGLCIAICAFTEFWIGDIIIVTYSPLVMLNAIGLAGVFSIYAVVCLISFVFVFLK 720
QY 774 VPETKGMPLLEVITEFFAVGAKQAAKA 800
DB 721 VPETKGMPLLEVITEFFAVGAKQAAKA 747

RESULT 4

US-10-051-909-2
; Sequence 2, Application US/10051909
; Publication No. US20020199217A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Helentjaris, Tim
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB1163 US CIP
; CURRENT APPLICATION NUMBER: US/10/051,909
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 60/083,044
; PRIOR FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 747
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (129)
; NAME/KEY: UNSURE
; LOCATION: (133)..(134)

```

; NAME/KEY: UNSURE
; LOCATION: (144)
; NAME/KEY: UNSURE
; LOCATION: (178)
; NAME/KEY: UNSURE
; LOCATION: (207)
; NAME/KEY: UNSURE
; LOCATION: (218)
; NAME/KEY: UNSURE
; LOCATION: (220)
; NAME/KEY: UNSURE
; LOCATION: (236)
US-10-051-909-2

Query Match      86.3%; Score 3517; DB 13; Length 747;
Best Local Similarity 91.6%; Pred. No. 1.4e-293;
Matches 684; Conservative 27; Mismatches 34; Indels 2; Gaps 2;

QY 56 MSGAVLVAIVASIGNLLQGDNDATIAAAVLYIKKEFQQLONEPTVEGLIVMSLIGATIVT 115
Db 1 MGVAVMAIAASIGNLLQGDNDATIAAGAVLYIKKEFNQLQSEPLIEGLIVAMFLIGATVIT 60

QY 116 TFSGPLSDSIGRRPMLILSSILYFFSGLMLWSPNVYVLLARFVDGFGIGLAVTLVPLY 175
Db 61 TSPGPRADCVGRPRMLVASAVLYFVSGLVMLWAPIVYILLARLIDGFGIGLAVTLVPLY 120

QY 176 ISEIAPSEIR-GLLNTLPQFSG-SGGMFLSYCMVFGMSLSPSPDWRLMLGLVALPSLFF 233
Db 121 ISETAPHRXSWGXXNTLPQFIGVXGGMFLSYCMVFGMSLMPKPDWRLMLGLVLSIPSLXYF 180

QY 234 GLTIFYPESPRWLVS KGRMAEAKKVLQKLRGKDDVSGELSLLLEGLVGGDTSEIEYII 293
Db 181 GLTVFYPESPRWLVS KGRMAEAKKRVXQRLRGREDVSXEXALLVEGLVGKDTRIXEYII 240

QY 294 GPATEAADLVTDGKEQITLYGPEEGQSWIARPSKGPIMLSVLSASRHGSMVNQSV 353
Db 241 GPATEAADLVTDGKEQITLYGPEEGQSWIARPSKGPIMLSVLSASRHGSMVNQSV 300

QY 354 LMDPIVTLFGSVHENMPQAGGMRSTLFPNFGSMFVSVDQHAQNEQWDEENLHRDDEYA 413
Db 301 LMDPIVTLFGSVHENMPQAGGMRSTLFPNFGSMFVSVDQHAQNEQWDEENLHRDDEYA 360

QY 414 SDGAGGDIYEDNLHSPLLSRQATGAEGKDIVHHGHRGSALS MRQTLLGEGGDSVSTDIG 473
Db 361 SDGAGGDIYEDNLHSPLLSRQATGAEGKDIVHHGHRGSALS MRQTLLGEGGDSVSTDIG 420

QY 474 GGWLAWKWEKEGNGRKEGGFKRVYLHQEGVPGSRGSI VSLPGGGDVFESEFVHAA 533
Db 421 GGWLAWKWEKEGNGRKEGGFKRVYLHQEGVPGSRGSI VSLPGGGDVFESEFVHAA 480

QY 534 ALVSQSALFSKGLAEPMSDAAVHPSEVAAGSRWKDLFEPGVRALLLVGVGIQLQQF 593
Db 481 ALVSQSALFSKGLAEPMSDAAVHPSEVAAGSRWKDLFEPGVRALLLVGVGIQLQQF 540

QY 594 AGINGVLYTTPQILEQAGVAVILSKFGLSSASASILISLTLMLPCIGFAMLLMDLSG 653
Db 541 AGINGVLYTTPQILEQAGVAVILSKFGLSSASASILISLTLMLPCIGFAMLLMDLSG 600

QY 654 RRFLLLGTIPILIASLVILVNSNLDLGTLAHALLSTVSIVVYFCCFVMGFGPIPNILCA 713
Db 601 RRFLLLGTIPILIASLVILVNSNLDLGTLAHALLSTVSIVVYFCCFVMGFGPIPNILCA 660

QY 714 EIFPTRVRGLCIAICAFTEFWIGDIIVTYSLPVMLNAIGLAGVFSIYAVVCLISFVFVFLK 773
Db 661 EIFPTRVRGLCIAICAFTEFWIGDIIVTYSLPVMLNAIGLAGVFSIYAVVCLISFVFVFLK 720

QY 774 VPETKGMPLVITEFFAVGAKQAAKA 800
Db 721 VPETKGMPLVITEFFAVGAKQAAKA 747
```

```

; Sequence 56898, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 56898
; LENGTH: 564
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLMO17045C09_FLI.ppep
US-10-425-114-56898

Query Match      70.4%; Score 2869; DB 15; Length 564;
Best Local Similarity 99.5%; Pred. No. 6.2e-238;
Matches 561; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 237 IFYLPESPRWLVS KGRMAEAKKVLQKLRGKDDVSGELSLLLEGLVGGDTSEIEYII 296
Db 1 IFYLPESPRWLVS KGRMAEAKKVLQKLRGKDDVSGELSLLLEGLVGGDTSEIEYII 60

QY 297 TEAADLVTDGKEQITLYGPEEGQSWIARPSKGPIMLSVLSASRHGSMVNQSVPLMD 356
Db 61 TEAADLVTDGKEQITLYGPEEGQSWIARPSKGPIMLSVLSASRHGSMVNQSVPLMD 120

QY 357 PIVTLFGSVHENMPQAGGMRSTLFPNFGSMFVSVDQHAQNEQWDEENLHRDDEYASDG 416
Db 121 PIVTLFGSVHENMPQAGGMRSTLFPNFGSMFVSVDQHAQNEQWDEENLHRDDEYASDG 180

QY 417 AGGDIYEDNLHSPLLSRQATGAEGKDIVHHGHRGSALS MRQTLLGEGGDSVSTDIGGW 476
Db 181 AGGDIYEDNLHSPLLSRQATGAEGKDIVHHGHRGSALS MRQTLLGEGGDSVSTDIGGW 240

QY 477 QLAWKWEKEGNGRKEGGFKRVYLHQEGVPGSRGSI VSLPGGGDVFESEFVHAAALV 536
Db 241 QLAWKWEKEGNGRKEGGFKRVYLHQEGVPGSRGSI VSLPGGGDVFESEFVHAAALV 300

QY 537 QSALFSKGLAEPMSDAAVHPSEVAAGSRWKDLFEPGVRALLLVGVGIQLQQFAGI 596
Db 301 QSALFSKGLAEPMSDAAVHPSEVAAGSRWKDLFEPGVRALLLVGVGIQLQQFAGI 360

QY 597 NGVLYTTPQILEQAGVAVILSKFGLSSASASILISLTLMLPCIGFAMLLMDLSGRFF 656
Db 361 NGVLYTTPQILEQAGVAVILSKFGLSSASASILISLTLMLPCIGFAMLLMDLSGRFF 420

QY 657 LLLGTIPILIASLVILVNSNLDLGTLAHALLSTVSIVVYFCCFVMGFGPIPNILCAIF 716
Db 421 LLLGTIPILIASLVILVNSNLDLGTLAHALLSTVSIVVYFCCFVMGFGPIPNILCAIF 480

QY 717 PTRVRGLCIAICAFTEFWIGDIIVTYSLPVMLNAIGLAGVFSIYAVVCLISFVFVFLK 776
Db 481 PTRVRGLCIAICAFTEFWIGDIIVTYSLPVMLNAIGLAGVFSIYAVVCLISFVFVFLK 540

QY 777 TKGMPLVITEFFAVGAKQAAKA 800
Db 541 TKGMPLVITEFFAVGAKQAAKA 564
```


APPLICANT: Allen, Steve
APPLICANT: Helentjaris, Tim
APPLICANT: Hitz, Bill
APPLICANT: Kinney, Tony
APPLICANT: Tingey, Scott
TITLE OF INVENTION: Plant Sugar Transport Proteins
FILE REFERENCE: BB1163 US CIP
CURRENT APPLICATION NUMBER: US/10/051,909
CURRENT FILING DATE: 2002-01-17
PRIOR APPLICATION NUMBER: 60/083,044
PRIOR FILING DATE: April 24, 1998
NUMBER OF SEQ ID NOS: 38
SOFTWARE: Microsoft Office 97
SEQ ID NO 37
LENGTH: 740
TYPE: PRT
ORGANISM: Oryza sativa
US-10-051-909-37

Query Match 69.5%; Score 2834; DB 13; Length 740;
Best Local Similarity 74.5%; Pred. No. 9.7e-235;
Matches 553; Conservative 75; Mismatches 108; Indels 6; Gaps 5;

QY 56 MSGAVLVAIVASIGNLLQGDWNTAAAVLYIKKEFQIQNEPTVEGLIVMSLIGATIVT 115
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
Db 1 MAGAVLVAIAASIGNLLQGDWNTAAAGAVLYIKKEFNQSEPLIEGLIVMSLIGATIIT 60
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|

QY 116 TFSGPLSDSIGRRPMLILSSILYFFSGLIMLWSPNVYVLLARFVDGFGIGLAVTLVPLY 175
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
Db 61 TFSGAVADSFGRPRMLIASAVLYFVSGLVMLWAPNVYVLLARLIDGFGIGLAVTLVPLY 120
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|

QY 176 ISEIAPSEIRGLLNTLPQFSGSGMFLSYCMVFGMSLSPSPDWRIMLGVLAIPLSFFFL 235
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
Db 121 ISETAPTDIRGLLNTLPQFSGSGMFLSYCMVFGMSLMPQPDWRIMLGVLISPLIYFAL 180
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|

QY 236 TIFYPESPRWLVS KGRMAEAKVLQKRGDDVSGELSILLEGLEGGDTSEIEYIIGP 295
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
Db 181 TIFYPESPRWLVS KGRMAEAKVLQGLRGREDVSGEMALLVEGLGVGKDTKIEEYIIGP 240
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|

QY 296 ATEAADLVTGDKEQITLYGPEGQSWIARPSKGPIMLGSVLSASRHGSMVNSQVPLM 355
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
Db 241 DDELADEGLAP-DPEKIKLYGPEGLSWARPVHGQSALGSLGLISRHGSMVSQKPLV 299
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|

QY 356 DPIVTLFGSVHENMPQAGGSMRSTLFPNFGSMFVSVDQHAQNEQWDEENLHRDDEEYASD 415
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
Db 300 DPVVTLFGSVHEKMP EIMGSMRSTLFPNFGSMFVSVAEQQAQKGDWDAES-QREGEDYGSD 358
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|

QY 416 GAGGDYEDNLHSPLLSRQATGAEGKDI-VHHGHRGSALSMMRQTLLGEGGDGVSSTDIGG 474
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
Db 359 HGGDDIEDLSQSP LISRQATSVGEKEIAAPHGSIIMGAVG--RSSSLMQGGEAVSSMGIGG 416
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|

QY 475 GWQLAWKWEKEGNGRKEGGFKRVYLHQEGVPGSRRGSI VSLPGGDDVFESEFVHAAA 534
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
Db 417 GWQLAWKWT EREGADGEKEGGFORIYLHEEGVTGDRRGSILSLP-GGDVPPGGEFVQAAA 475
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|

QY 535 LVSQSALFSKGLAEPRMSDAAVMHPSEVAAGSRWKDLFEPGVRALLVGVGIQILQQFA 594
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
Db 476 LVSQPALYSKELMEQRLAGPAMVHP SQAVAKGPKWADLFEPGVKHALFVGIGIILQQFA 535
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|

QY 595 GINGVLVYTPQILEQAGVAVILSKFGLSSASASILISLTLTLLMPCIGFAMLLMDLSGR 654
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
Db 536 GINGVLVYTPQILEQAGVGVLLANIGLSSSSASILISGLTLLMPSIGIAMRLMDMSGR 595
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|

QY 655 RFLLLGTPILIASLVILVVSNLIDGLTALHALLSTVSVIYVFCFVMGFGPIPNILCAE 714
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
Db 596 RFLLLATIPILIVALAILVNLIDVGTWVHASLSTVSVILYFCFVVMGFGPIPNILCAE 655
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|

QY 715 IFPTRVGLICIAICFTFWIGDIIYVYSLPVMNLNAIGLAGVFSIYAVVCLISFVFVFLKV 774
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
Db 656 IFPTTVGLICIAICALTFWIGDIIYVYTL PVMNLNAIGLAGVGIYAVVCILAFLFVFMKV 715
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|

QY 775 PETKGMPLLEVITEFFAVGAKQA 796
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|

Db 716 PETKGMPLLEVITEFFSVGAKQA 737

RESULT 7
US-10-437-963-195955
Sequence 195955, Application US/10437963
Publication No. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 195955
LENGTH: 740
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_91853C.1.pep
US-10-437-963-195955

Query Match 69.5%; Score 2834; DB 16; Length 740;
Best Local Similarity 74.5%; Pred. No. 9.7e-235;
Matches 553; Conservative 75; Mismatches 108; Indels 6; Gaps 5;

QY 56 MSGAVLVAIVASIGNLLQGDWNTAAAVLYIKKEFQIQNEPTVEGLIVMSLIGATIVT 115
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
Db 1 MAGAVLVAIAASIGNLLQGDWNTAAAGAVLYIKKEFNQSEPLIEGLIVMSLIGATIIT 60
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|

QY 116 TFSGPLSDSIGRRPMLILSSILYFFSGLIMLWSPNVYVLLARFVDGFGIGLAVTLVPLY 175
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
Db 61 TFSGAVADSFGRPRMLIASAVLYFVSGLVMLWAPNVYVLLARLIDGFGIGLAVTLVPLY 120
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|

QY 176 ISEIAPSEIRGLLNTLPQFSGSGMFLSYCMVFGMSLSPSPDWRIMLGVLAIPLSFFFL 235
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
Db 121 ISETAPTDIRGLLNTLPQFSGSGMFLSYCMVFGMSLMPQPDWRIMLGVLISPLIYFAL 180
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|

QY 236 TIFYPESPRWLVS KGRMAEAKVLQKRGDDVSGELSILLEGLEGGDTSEIEYIIGP 295
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
Db 181 TIFYPESPRWLVS KGRMAEAKVLQGLRGREDVSGEMALLVEGLGVGKDTKIEEYIIGP 240
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|

QY 296 ATEAADLVTGDKEQITLYGPEGQSWIARPSKGPIMLGSVLSASRHGSMVNSQVPLM 355
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
Db 241 DDELADEGLAP-DPEKIKLYGPEGLSWARPVHGQSALGSLGLISRHGSMVSQKPLV 299
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|

QY 356 DPIVTLFGSVHENMPQAGGSMRSTLFPNFGSMFVSVDQHAQNEQWDEENLHRDDEEYASD 415
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
Db 300 DPVVTLFGSVHEKMP EIMGSMRSTLFPNFGSMFVSVAEQQAQKGDWDAES-QREGEDYGSD 358
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|

QY 416 GAGGDYEDNLHSPLLSRQATGAEGKDI-VHHGHRGSALSMMRQTLLGEGGDGVSSTDIGG 474
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
Db 359 HGGDDIEDLSQSP LISRQATSVGEKEIAAPHGSIIMGAVG--RSSSLMQGGEAVSSMGIGG 416
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|

QY 475 GWQLAWKWEKEGNGRKEGGFKRVYLHQEGVPGSRRGSI VSLPGGDDVFESEFVHAAA 534
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
Db 417 GWQLAWKWT EREGADGEKEGGFORIYLHEEGVTGDRRGSILSLP-GGDVPPGGEFVQAAA 475
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|

QY 535 LVSQSALFSKGLAEPRMSDAAVMHPSEVAAGSRWKDLFEPGVRALLVGVGIQILQQFA 594
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
Db 476 LVSQPALYSKELMEQRLAGPAMVHP SQAVAKGPKWADLFEPGVKHALFVGIGIILQQFA 535
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|

QY 595 GINGVLVYTPQILEQAGVAVILSKFGLSSASASILISLTLTLLMPCIGFAMLLMDLSGR 654
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
Db 536 GINGVLVYTPQILEQAGVGVLLANIGLSSSSASILISGLTLLMPSIGIAMRLMDMSGR 595
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|

QY 655 RFLLLGTIPILIASLVILVSNLIDLGLTIAHALLSTVSVIVYFCCFVMGFGPIPNILCAE 714
Db 596 RFLLLATIPILIVALAILLVNILDVGTVMHASTVSVILYFCFFVMGFGPIPNILCAE 655
QY 715 IFPTRVRGLCIAICAFTEFWIGDIIIVTYSLPVMLNAIGLAGVFSIYAVVCLISFVFELKV 774
Db 656 IFPTTVRGICIAICALTFWIGDIIIVTYPVMLNAIGLAGVFGIYAVVCIILAFLEFVFMKV 715
QY 775 PETKGMPLVITEFFAVGAKQA 796
Db 716 PETKGMPLVITEFFSVGAKQA 737
RESULT 8
US-10-051-902-8
; Sequence 8, Application US/10051902
; Publication No. US20020178468A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/10/051,902
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-14
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 8
; LENGTH: 737
; TYPE: PRT
; ORGANISM: Glycine max
US-10-051-902-8

Query Match 65.6%; Score 2674; DB 13; Length 737;
Best Local Similarity 69.9%; Pred. No. 5.9e-221;
Matches 522; Conservative 92; Mismatches 115; Indels 18; Gaps 8;

QY 56 MSGAVLVAIVASIGNLLQGDNDATIAAAVLYIKKEFQLQNEPTVEGLIVSMSLIGATVIT 115
Db 1 MKGAVLVAIAASIGNFLQGDNDATIAAGANGYIKKDLALGT--TMERLVVGMSLIGATVIT 58
QY 116 TFSGPLSDSICRRPMLILSSILYFFSGGGMFLSYCMVFGMSLSPSPDWRIMLGVLAIPLSFFFG 175
Db 59 TCSGPIADWLGRPPMMIISVLYFLGLVLMWSPNVYVLCARLLDGGFGLAVTLVPVY 118
QY 176 ISEIAPSEIRGLLNTLPQFSGSGGGMFLSYCMVFGMSLSPSPDWRIMLGVLAIPLSFFFG 235
Db 119 ISETAPSEIRGLSNTLPQFSGSGGGMFLSYCMVFGMSLSPAPSWRLMLGVLSIPSLLYFAL 178
QY 236 TIFYPESPRWLVSIGRMMAEAKKVLQKLRGKDDVSGELSLLEGLVGGDTSIIEYIIGP 295
Db 179 TIFFLPESPRWLVSIGRMMAEAKKVLQRLRGREDVSGEMALLVEGLIGGDTSIIEYIIGP 238
QY 296 ATEAADLVTDGDKQITLYGPEEGQSWIARPSKGPIMLGSVLSASRHGSMVNSQSVPLM 355
Db 239 ADDVADGHEHATEKDRIKIRLYGSQAGLSWLSKPVGTQSSIG----LASHHGSIINQSMPLM 294
QY 356 DPIVTLFGSVHENMPQAG--GSMRSTLFPNFGSMFVSVDQHAKEQWDEENLHRDDEEYA 413
Db 295 DPLVTLFGSIHEKLPETGARGSMRSTLFPNFGSMFSTAEPHAKIEQWDEESLQREDEDYM 354
QY 414 SDGAGGDYEDNLHSPLLSRQATGAEGKDIVHHGHRGSAL-SMRQTLLGEG-GDGVSTTD 471
Db 355 SDATRGSDDDNLHSPILSRQTTSL-KDLPPPPSHGSIILGSMRHSLSMQSGEGQGSGTG 413
QY 472 IGGWQLAWKWSKEGEGNKGKRGKRVYLHQGVPGSRRGSIYSLPGGGDVFESEFVH 531
Db 414 IGGWQLAWKWTDK-GEDGKQGGFKRIYLHEEGVSASRRGSTVSIPEG-----BFRVQ 466

QY 532 AAALVVSQALFSKGLAEPRMSDAAMVHPSEVAAGSRWKDLFEPGVRRAALLVGVGIQILQ 591
Db 467 AAALVVSQALYSKELIDGHPVGPAMVHPSETASKGSPSWKALLEPGVKHALVVGVIQILQ 526
QY 592 QFAGINGVLYYTPQILEQAGVAVILSKFGISSASASILISSLTLLMLPCIGFAMLLMDL 651
Db 527 QFSGINGVLYYTPQILEAGVEVLLSDIGIGSESASFLLISAFITFLMLPCIGVAMKLMV 586
QY 652 SGRFLLLTIPILIASLVILVSNLIDLGLTIAHALLSTVSVIVYFCCFVMGFGPIPNIL 711
Db 587 SGRQLLLTIPVLIVSLIILVIGSLVNFNGVAHAISTVCVVVYFCCFVMGFGPIPNIL 646
QY 712 CAEIFPTRVRGLCIAICAFTEFWIGDIIIVTYSLPVMLNAIGLAGVFSIYAVVCLISFVFE 771
Db 647 CSEIFPTRVRGLCIAICALTFWIGDIIIVTYPVMLGSLGLGVFAIYAVVCFISWIFV 706
QY 772 LKVPETKGMPLVITEFFAVGAKQA 798
Db 707 LKVPETKGMPLVITEFFSVGAKQAAS 733

RESULT 9
US-10-051-909-8
; Sequence 8, Application US/10051909
; Publication No. US20020199217A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Helentjaris, Tim
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB1163 US CIP
; CURRENT APPLICATION NUMBER: US/10/051,909
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 60/083,044
; PRIOR FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 8
; LENGTH: 737
; TYPE: PRT
; ORGANISM: Glycine max
US-10-051-909-8

Query Match 65.6%; Score 2674; DB 13; Length 737;
Best Local Similarity 69.9%; Pred. No. 5.9e-221;
Matches 522; Conservative 92; Mismatches 115; Indels 18; Gaps 8;

QY 56 MSGAVLVAIVASIGNLLQGDNDATIAAAVLYIKKEFQLQNEPTVEGLIVSMSLIGATVIT 115
Db 1 MKGAVLVAIAASIGNFLQGDNDATIAAGANGYIKKDLALGT--TMERLVVGMSLIGATVIT 58
QY 116 TFSGPLSDSICRRPMLILSSILYFFSGGGMFLSYCMVFGMSLSPSPDWRIMLGVLAIPLSFFFG 175
Db 59 TCSGPIADWLGRPPMMIISVLYFLGLVLMWSPNVYVLCARLLDGGFGLAVTLVPVY 118
QY 176 ISEIAPSEIRGLLNTLPQFSGSGGGMFLSYCMVFGMSLSPSPDWRIMLGVLAIPLSFFFG 235
Db 119 ISETAPSEIRGLSNTLPQFSGSGGGMFLSYCMVFGMSLSPAPSWRLMLGVLSIPSLLYFAL 178
QY 236 TIFYPESPRWLVSIGRMMAEAKKVLQKLRGKDDVSGELSLLEGLVGGDTSIIEYIIGP 295
Db 179 TIFFLPESPRWLVSIGRMMAEAKKVLQRLRGREDVSGEMALLVEGLIGGDTSIIEYIIGP 238
QY 296 ATEAADLVTDGDKQITLYGPEEGQSWIARPSKGPIMLGSVLSASRHGSMVNSQSVPLM 355
Db 239 ADDVADGHEHATEKDRIKIRLYGSQAGLSWLSKPVGTQSSIG----LASHHGSIINQSMPLM 294
QY 356 DPIVTLFGSVHENMPQAG--GSMRSTLFPNFGSMFVSVDQHAKEQWDEENLHRDDEEYA 413
Db 295 DPLVTLFGSIHEKLPETGARGSMRSTLFPNFGSMFSTAEPHAKIEQWDEESLQREDEDYM 354

QY 414 SDGAGG DYEDNLHSP LLSRQATGAEGKDI VHHGHRGSAL-SMRRQTLLGEG-GDGVSTSD 471
Db 355 SDATRGDSDDLHSP LLSRQTTSLE-KDLPPPPSHGSI LGMRRHSSLMQSGSGEQGGSTG 413
QY 472 IGGWQLAWKSWEXKEGNGRKEGGFKRVYLHOBGVPGSRRGSI VSLPGGGDVFESEFVH 531
Db 414 IGGWQLAWKWTDK-CEDGKQGGFKRIYLHBEGVSA SRRGSI VSIPEG-----EFVQ 466
QY 532 AAALVSQSALFSKGLAEP RMSDAAMVHPSEVA AKGSRWKDLFEPGVRALLVGVGIQILQ 591
Db 467 AAALVSQPALYSKELLIDGHPVGPAMVHPSE TASKGPSWKALLEPGVKHALVVGVIQILQ 526
QY 592 QFAGINGVLYYTPQIL EQAGVAVILSKFGLSSASASILISSLTTLMLPCIGFAMLLMDL 651
Db 527 QFSGINGVLYYTPQIL EEAGVEVLLSDIGIGSESASF LISAF TTFLMLPCIGVAMKMDV 586
QY 652 SCRRFLLGTIPILIASLIVLVSNLIDLGT LAHALLSTVSIVIVFCCFVMGFGPIPNIL 711
Db 587 SCRRQLLLTTIPVLIVSLIILVIGSLVNFGNVAHA AISTVCVVVYFCCFVMGYGPIPNIL 646
QY 712 CAEIFPTRVRGLCIACAFTFWIGDIIIVTYSLPVMLNAI GLAGVFSIYAVVCLISFVFVF 771
Db 647 CSEIFPTRVRGLCIAICALVFWIGDIIITYSLPVMLGSLGLGGVFAIYAVVCFISWIFVF 706
QY 772 LKVPETKGMPL EVITEFFAVGAKQAAA 798
Db 707 LKVPETKGMPL EVISEFFSVGAKQAAS 733

RESULT 10

US-10-051-902-29
; Sequence 29, Application US/10051902
; Publication No. US20020178468A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/10/051,902
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-14
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 29
; LENGTH: 729
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-051-902-29

Query Match 63.4%; Score 2584; DB 13; Length 729;
Best Local Similarity 68.8%; Pred. No. 3.3e-213;
Matches 512; Conservative 89; Mismatches 123; Indels 20; Gaps 8;

QY 56 MSGAVLVAIVASIGNLLQGW DNATIAAAVLYIKKEFQLQNEPTVEGLIVMSLIGATIVT 115
Db 1 MSGAVLVAIAAAVGNLLQGW DNATIAAGAVLYIKKEFNLESNPSVEGLIVAMSLIGATLIT 60
QY 116 TFSGPLSDSIGRRPMLILSSILYFFSG LIMLWSPNVYVLLARFVDGFGIGLAVTLVPLY 175
Db 61 TCSGGVADWLGRRPMLILSSILYFVGS LVMWSPNVYVLLGRLLDGFVGLVTLVPIY 120
QY 176 ISEIAPSEIRGLLNTLPQFSGGGMFLSYCMVFGMSLSPDPWRIMLGVLAI PSLFFFG 235
Db 121 ISETAPEIRGLLNTLPQFTSGGGMFLSYCMVFGMSLMPSPWRMLMLGVLPFIPSLVFFFL 180
QY 236 TIFYPESPRLVSKGRMAEAKVQLKRGKDDVSGELSLLEGLVGGDTSI EYIIGP 295
Db 181 TVFFLPESPRLVSKGRMLEAKRVLQRLRGREDVSGEMALLVEGLIGGETTIEEYIIGP 240
QY 296 ATEAADLVTDGDK EQITLYGPEEGQSWIARPSKGPIMLGSVLSASRHGS-MV NQSVPL 354

Db 241 ADEVTDHDDIAVDKDI KLYGAEEGLSWARPVKG-----GSTMSVL SRHGSTMRRQGSL 296
QY 355 MDPIVTLFGSVHENMPQAGGSMRSTLFPNFGSMF SVTDQHAKNEQWDEENLHRDDEEYAS 414
Db 297 IDPLVTLFGSVHEKMPDT-GSMRSALFPHFGSMF SVGGNQPRHEDWDENLVGEGEDYPS 355
QY 415 DGAGGDYEDNLHSP LLSRQATGAEGKDI VHHGHRGSALSMMRQTLLGEGGDVGSSTDIGG 474
Db 356 D-HGDDSEDDLHSP LLSRQTTSME-KDMPHTAHGTLSTFRHGSQVQGAQGE GAGSMGIGG 413
QY 475 GWQLAWKSWEXKEGNGRKEGGFKRVYLHQBGVPGSRRGSI VSLPGGGDVFESEFVHAAA 534
Db 414 GWQVANKWTEREDESQKE-----EGFPGRRGSI VSLPGGGDTGE-AD FVQAASA 462
QY 535 LVSQSALFSKGLAEP RMSDAAMVHPSEVA AKGSRWKDLFEPGVRRALLVGVGIQILQQFA 594
Db 463 LVSQPALYSKDLLKHEHTIGFAMVHPSE-TTKGSIWHD LHDHPGVKRALVVGVLQILQQFS 521
QY 595 GINGVLYYTPQIL EQAGVAVILSKFGLSSASASILISSLTTLMLPCIGFAMLLMDLSGR 654
Db 522 GINGVLYYTPQIL EQAGVGVGILLSNMGISSSSASLLISALTTFVMLPAI AVAMRLMDLSGR 581
QY 655 RFLLLGTIPILIASLIVLVSNLIDLGT LAHALLSTVSIVIVFCCFVMGFGPIPNILCAE 714
Db 582 RTLLTTIPILIASLIVLVISNLVHMNSIVH AVLSTVSVVLYFCFFVMGFGPAPNILCSE 641
QY 715 IFPTRVRGLCIACAFTFWIGDIIIVTYSLPVMLNAI GLAGVFSIYAVVCLISFVFVLKV 774
Db 642 IFPTRVRGICIAICALTFWICDIIIVTYSLPVLLKSI GLAGVFGMYAIVCCISWVVFVKV 701
QY 775 PETKGMPL EVITEFFAVGAKQAAA 798
Db 702 PETKGMPL EVITEFFSVGARQAEA 725

RESULT 11

US-10-051-909-29
; Sequence 29, Application US/10051909
; Publication No. US20020199217A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Helentjaris, Tim
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB1163 US CIP
; CURRENT APPLICATION NUMBER: US/10/051,909
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 60/083,044
; PRIOR FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 29
; LENGTH: 729
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-051-909-29

Query Match 63.4%; Score 2584; DB 13; Length 729;
Best Local Similarity 68.8%; Pred. No. 3.3e-213;
Matches 512; Conservative 89; Mismatches 123; Indels 20; Gaps 8;

QY 56 MSGAVLVAIVASIGNLLQGW DNATIAAAVLYIKKEFQLQNEPTVEGLIVMSLIGATIVT 115
Db 1 MSGAVLVAIAAAVGNLLQGW DNATIAAGAVLYIKKEFNLESNPSVEGLIVAMSLIGATLIT 60
QY 116 TFSGPLSDSIGRRPMLILSSILYFFSG LIMLWSPNVYVLLARFVDGFGIGLAVTLVPLY 175
Db 61 TCSGGVADWLGRRPMLILSSILYFVGS LVMWSPNVYVLLGRLLDGFVGLVTLVPIY 120
QY 176 ISEIAPSEIRGLLNTLPQFSGGGMFLSYCMVFGMSLSPDPWRIMLGVLAI PSLFFFG 235

Db 121 ISETAPPEIRGLNLTLPQFTGSGGMFLSYCMVFGMSLMPSPSWRLMLGVLPFIPSLVFFFL 180
QY 236 TIFYLPESPRWLVSCKGRMAEAKVQLKRLGKDDVSGELSLLLLEGVEGDDTSIEEYIIGP 295
Db 181 TVFFLPESPRWLVSCKGRMLEAKRVLQRLRGREDVSGEMALLVEGLGIGGETTIEEYIIGP 240
QY 296 ATEAADLVTGDKKEQITLYGPBEGQSWIARPSKGPIMLGSLVSLASRHGS-MVNQSVPL 354
Db 241 ADEVTDHDIADVQDIKLYGAEGLSWARPVKG---GSTMSVLSRHGSTMRRQGS 296
QY 355 MDPIVTLFGSVHENMPQAGGSMRSTLFPNFGSMFSDTDQHAKEQWDEENLHRDDEEYAS 414
Db 297 IDPLVTLFGSVHEKMPDT-GSMRSALFPFHFGSMFVSGNQPRHEDWDEENLVGEGEDYPS 355
QY 415 DGAGGDYEDNLHSLSRQATGAEGKDIVHGHGRGSALSMMRQTLLGEGDGVSTDIG 474
Db 356 D-HGDDSEDDLHSLPSRQTTSME-KDMPHTAHTLSTFRHGSQVQGAQGEAGSMGIG 413
QY 475 GWQLAWKSEKEGNGRKEGGFKRVYLHQEGVPGSRRGSIIVSLPGGDDVFEGSEFVHAAA 534
Db 414 GWQVAKWTEREDSGQKE-----EGPPGSRGSIIVSLPGGDDTGE-ADFFVQASA 462
QY 535 LVSQALFSKGLAEPRMSDAAMVHPSEVAAGSRWKDLFEPGVRRALLVGVGIQILOQFA 594
Db 463 LVSQPALYSKDLKEHTIGPAMVHPSE-TTKGSIWHDLHPGVKRALVGVGLQILOQFS 521
QY 595 GINGVLYTTPQILEQAGVAVILSKFGLSSASASILISLTLLMLPCIGFAMLLMDLSGR 654
Db 522 GINGVLYTTPQILEQAGVAVILSKFGLSSASASILISLTLLMLPCIGFAMLLMDLSGR 581
QY 655 RFLLLGTIPILIASLVILVSNLIDLTGLAHALLSTVSIVIVFCCFVMGFGPIPNILCAE 714
Db 582 RTLLLTTPILIASLVILVSNLIDLTGLAHALLSTVSIVIVFCCFVMGFGPIPNILCAE 641
QY 715 IFPTRVRGLCIACTFWIGDIIIVTYSIPVMLNAIGLAGVFSIYAVVCLISFVFVFLKV 774
Db 642 IFPTRVRGLCIACTFWIGDIIIVTYSIPVMLNAIGLAGVFSIYAVVCLISFVFVFLKV 701
QY 775 PETKGMPLVITTEFFAVGAKQAAA 798
Db 702 PETKGMPLVITTEFFSVGARQAEA 725

RESULT 12
US-10-424-599-153166
; Sequence 153166, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J.
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 153166
; LENGTH: 742
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(742)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_109333C.1.pap
US-10-424-599-153166
Query Match 63.3%; Score 2581.5; DB 15; Length 742;
Best Local Similarity 67.3%; Pred. No. 5.5e-213;

Matches 511; Conservative 101; Mismatches 110; Indels 37; Gaps 13;
QY 56 MSGAVLVAIVASIGNLLQGDNDATIAAAVLYIKKEFQIQNEPTVEGLIVSMSLIGATIVT 115
Db 1 MKGTVLVAIAAIGNILQGDNDASIAIAVYIKKDLALQT--TMEGLVAVMSLIGATVIT 58
QY 116 TFSGLSDSIGRRPMLILSSILYFFSGLIMLWSPNVYVLLARFVDCFGIGLAVTLVPLY 175
Db 59 TCSGPIADWLGRRRPMIISVLYFLGGLVMLWSPNVVLCARLLDGFGLAVTLVPVY 118
QY 176 ISEIAPSEIRGLNLTLPQFSGSGGMFLSYCMVFGMSLSPSPDWIRIMLGVLAIPLSFFFL 235
Db 119 ISETAPSEIRGSLNLTLPQVSGSGGMFLSYFMVFGMSLSPAPSWRLMLGVLSIPSLLYFAL 178
QY 236 TIFYLPESPRWLVSCKGRMAEAKVQL-----QKLRGKDDVSGELSLLLLEG-LEVGGDTS 287
Db 179 TIFYLPES-----SSGSVSKXNLMRLRQTKCSQRLRGREDVSGEMTLLVEGSTXIGDTS 234
QY 288 IEEYIIGPATEAADLVTGDKKEQITLYGPBEGQSWIARPSKGPIMLGSLVSLASRHGSM 347
Db 235 IEEYIIGXADQVADGHEHATEKDRIKLYGSAQAGLSWLAKPVTGQSSIG----LASRHGSI 290
QY 348 VNQSVPLMDPIVTLFGSVHENMPQ--AGGSMRSTLFPNFGSMFSDTDQHAKEQWDEENL 405
Db 291 INQSMPLMDPLVTLFGSIHEKLPETGAGGSMRSTLFPNFGSMFSTAEPAHAKNEQWDEESL 350
QY 406 HRDDEYASDGAGDYEDNLHSLSRQATGAEGKDI---VHGHGRGSAL-SMRRQTLL 460
Db 351 QREGEDYMSDAAGGDDSDNLHSLPSRQTTSLE-KDLPPPPSH--GSILGSMRRHSSL 406
QY 461 GEG-GDGVSTDIGGWQLAWKSEKEGNGRKEGGFKRVYLHQEGVPGSRRGSIIVSLPG 519
Db 407 MQGSGEQGGTGIGGWQLAWKWTDX-GEDGKQGGFKRIYLHEEGVSAASRRGSIIVSIPG 465
QY 520 GGDVFESEFVHAAALVSQSALFSKGLAEPRMSDAAMVHPSEVAAGSRWKDLFEPGVRR 579
Db 466 EG-----EFVQAAALVSQPALYSKELIDGHPVGPAMVHPSETASKGPSWKALLEPGVKH 519
QY 580 ALLVGVIQIILQOFAGINGVLYTTPQILEQAGVAVILSKFGLSSASASILISLTLLML 639
Db 520 ALVVGVIQIILQOFSGINGVLYTTPQILEEAGVEVLSDIGIGSESASFLLISAFITFLML 579
QY 640 PCIGFAMLLMDLSGRRFLLLGTIPILIASLVILVSNLIDLTGLAHALLSTVSIVIVFCC 699
Db 580 PCIGVAMKLMVDVSGRRQLLLTTPVLIGSLIILVIGSLVNFVNAHAAISTVCVVVYFCC 639
QY 700 FVMGFGPIPNILCAEIPTRVRGLCIACTFWIGDIIIVTYSIPVMLNAIGLAGVFSIY 759
Db 640 FVMGFGPIPNILCAEIPTRVRGLCIACTFWIGDIIIVTYSIPVMLNAIGLAGVFSIY 699
QY 760 AVCLISFVFVFLKVPEKGMPLVITTEFFAVGAKQAAA 798
Db 700 AVVCFISWIFVFLKVPEKGMPLVITTEFFAVGAKQAAA 738

RESULT 13
US-10-425-114-47258
; Sequence 47258, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 47258


```

; LENGTH: 728
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700728773_FLI.pep
US-10-425-114-47258

Query Match      62.4%; Score 2543; DB 15; Length 728;
Best Local Similarity 68.4%; Pred. No. 1.1e-209;
Matches 508; Conservative 83; Mismatches 120; Indels 32; Gaps 10;

QY 56 MSGAVLVAIVASIGNLLQGDWNTAIAAAVLYIKKEFQLQNEPTVEGLIVSMSLIGATIVT 115
Dd 10 MSGAVLVAVAAAGNLLQGDWNTAIAAGSILYIKREFNLQSEPTIEGLIVAMSLIGATVVT 69

QY 116 TFSGPLSDSICRRPMLILSSILYFFSGLIMLWSPNVYVLLARFVDFGIGLAVTLVPLY 175
Dd 70 TCSGPLSDLLGRRPMLIISSILYFVGSVLWMLWSPNVYILLFARLLDGLGIGLAVTLVPLY 129

QY 176 ISEIAPSEIRGLNLTLPQFSGSGMFLSYCMVFGMSLSPSPDWIRIMLGVLAIPLSFFFL 235
Dd 130 ISETAPPEIRGLNLTLPQFTGSAGMFFSYCMVFAISLTAKPNWRMLGLVLSIPSLIYFAL 189

QY 236 TIFYLPESPRLVSKGRMAEAKKVLQKLRGKDDYSGELSLLEGLVGGDTISIEEYIIGP 295
Dd 190 TLFFLPESPRLVSKGRMLEAKKVLQRLRGRQDVAGEMALLVEGLVGGDTVIEEYIISP 249

QY 296 ATEAADDLVTGDKEQITLYGPEEGQSWIARPSKGPIMLGSVLSASRHGSMVNQSVPLM 355
Dd 250 ANEFS-----DAEQIKLYGTAEQQSWIAKPVGTQSSIG---LVSRKGSMAHQSA-LV 297

QY 356 DPIVTLFGSVHENMPQAGGSMRSTLFPNFGSMFSDTDQAKNEQWDEENLHRDDEEYASD 415
Dd 298 DPLVKLFGSVHEKLPETG----STLFFPHFGSMFSGVGNQPRNEDWDEESIAREGDDYVSD 353

QY 416 GAG-GDYEDNLHSPILSRQATGAEGKDIVHHGHRGSALSMRQTLLEGGDGVSSTDIG 474
Dd 354 AADTDDSDNLQSPILSRQATSAE-RDMPAPAQG---SMRQGSLL--QGEPAQNSGIGG 406

QY 475 GWQLAWKWKSEKEGENGRKEGGFKRVYLHQEGVPGSRRGSIVSLPG-GGDVFEGSEFVHAA 533
Dd 407 GWQLAWKWKSETEGV-----FKRIYLHQEGGPGSRRGSIISLPGCDAPLTLDGEIVQAA 459

QY 534 ALVSQSALFSKGLAEPRMSDAAMVHPSEVAAKGSRWKDLFEPGVRRLALLVGVIQILQQF 593
Dd 460 ALVSQSALYNKELMHQQPVGPAMIHPSPQTAAKGPSWSDLLEPGVKHALIVGVGIQILQQF 519

QY 594 AGINGVLYYTPQILEQAGVAVILSKFGLSSASASILISSLTLLMLPCIGFAMLLMDLSG 653
Dd 520 SGINGVLYYTPQILEQAGVGVLLSNLGLGSTASFLISSVTLLMLPCIAVAMRLMDISG 579

QY 654 RPFLLGTIPILIASLVILVVSNLIDLGTLAHALLSTVSVIVYFCFVMGFGPIPNILCA 713
Dd 580 RPPLLTTPVLIIVSLLILVIGSLVELDSTINAFISTSSVIVYFCFVMGFGPIPNILCS 639

QY 714 EIPPTRVRGLCIAICAFTEFWIGDIIIVTYSPLPVMNLNAGLAGVFSIYAVVCLISFVFVLK 773
Dd 640 EIPPTRVRGLCIAICALTFWICDIIIVTYSPLPVMNSVGLGGVFGMYAVVVCIIANVFVFLK 699

QY 774 VPETKGMPLVITEFFAVGAKQA 796
Dd 700 VPETKGMPLVIIIEFFSVGAKQA 722
```

RESULT 14

```

US-10-424-599-251277
; Sequence 251277, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
```

```

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 251277
; LENGTH: 742
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(742)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_68931C.1.pep
US-10-424-599-251277
```

Query Match 61.3%; Score 2496.5; DB 15; Length 742;
Best Local Similarity 65.8%; Pred. No. 1.2e-205;
Matches 505; Conservative 84; Mismatches 121; Indels 57; Gaps 12;

```

QY 56 MSGAVLVAIVASIGNLLQGDWNTA-----AAVLYIKKEP 91
Dd 1 MSGAVLVAVAAAGNLLQGDWNTAIAGHIFIRLANXNDNIQPSLIGFLCYAGSILYIKREF 60

QY 92 QLQNEPTVEGLIVSMSLIGATIVTTFSGPLSDSISGRPMLILSSILYFFSGLIMLWSPNV 151
Dd 61 NLQSEPTIEGLIVAMSLIGATVVTTCGPLSDLLGRRPMLIISSILYFVGSVLWMLWSPNV 120

QY 152 YVLLLARFVDFGIGLAVTLVPLYIYSEIAPSEIRGLNLTLPQFSGSGMFLSYCMVFGMS 211
Dd 121 YILLFARLLDGLGIGLAVTLVPLYISETAPPEIRGLNLTLPQFTGSAGMFFSYCMVFAMS 180

QY 212 LSPSPDWIRIMLGVLAIPLSPFFFLGTIFYLPESPRWLVSIGRMAEAKKVLQKLRGKDDVSG 271
Dd 181 LTXAPNWRMLGLVLSIPSLIYFALTFLFPESPRWLV-KGRMLEAKKVLQRLRGRQDVAG 239

QY 272 ELSLLEGLVGGDTISIEEYIIGPATEAADDLVTGDKEQITLYGPEEGQSWIARPSKGP 331
Dd 240 EMALLVEGLVGRDTAIEEYIIGPAXEFS-----BAEQIKLYGTAEQSVWIAKPVGTQ 292

QY 332 IMLGSVLSLASRHGSMVNQSVPLMDPIVTLFGSVHENMPQAGGSMRSTLFPNFGSMFVS 391
Dd 293 SSIG---LVSRKGSMAHQSG-GLVDPVLKLFSGVHEKLPETG---STLFFPHFGSMFVS 343

QY 392 DQAKNEQWDEENLHRDDEEYASDQAG-GDYEDNLHSPILSRQATGAEGKDIVHHGHRGS 450
Dd 344 GNQPNEDWDEESIAREGDDYSDAVDTDDSDNLQSPILSRQATSAE-RDMPAPAQG-- 400

QY 451 ALSMERQTLLEGGDGVSSTDIGGWQLAWKWKSEKEGENGRKEGGFKRVYLHQEGVPGSR 510
Dd 401 --SMRQGSLL--QGEPAQNSGIGGWQLAWKWKSETEGV-----FKRIYLHQEGGPGSR 449

QY 511 RGSIVSLPG-GGDVFEGSEFVHAAALVSQSALFSKGLAEPRMSDAAMVHPSEVAAKGSRW 569
Dd 450 RGSIIISLPGCDAPLTLDGEIVQAAALVSQSALYNKELMHQQPVGPAMIHPSPQTAAGPSW 509

QY 570 KOLFEPGVRRLALLVGVIQILQQFAGINGVLYYTPQILEQAGVAVILSKFGLSSASASIL 629
Dd 510 SDLLEPGVKHALIVGVGIQILQQFSGINGVLYYTPQILEQAGVGVLLSNLGLGSTASFL 569

QY 630 ISSLTLLMLPCIGFAMLLMDISGRRFLLLTGTIPILIASLVILVVSNLIDLGTLAHALLS 689
Dd 570 ISSVTLLMLPCIAVAMRLMDISGRRTLLLTTPVLIVSLLILVIGSLVELDSTINAFIS 629

QY 690 TVSVIVYFCFVMGFGPIPNILCAEIPPTRVRGLCIAICAFTEFWIGDIIIVTYSPLPVM 749
Dd 630 TSSVIVYFCFVMGFGPIPNILCSEIFPTRVRGLCIAICALTFWICDIIIVTYSPLPVM 689

QY 750 IGLAGVFSIYAVVCLISFVFVFLKVPETKGMPLVITEFFAVGAKQA 796
Dd 690 VGLGGVFGMYAVVVCIIAWVFVFLKVPETKGMPLVIIIEFFSVGAKQA 736
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
OM protein - nucleic search, using frame_plus_p2n model
Run on: October 13, 2004, 11:39:29 ; Search time 9234 Seconds
(without alignments)
4097.007 Million cell updates/sec

Title: US-10-051-909-32
Perfect score: 4075
Sequence: 1 IRSGSWLAVQTFPTPLDRR.....PLEVITEFFAVGAKQAAKA 800

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4526729 seqs, 23644849745 residues
Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2 1/USPTO spool h/US10051909/runat 13102004 123336 19906/app_query.fasta_1.1678
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10051909@CGN 1 1 9091@runat 13102004 123336 19906 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*
1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES					
Result No.	Score	Query Match	Length	ID	Description
1	3517	86.3	2824	6 AR208565	AR208565 Sequence
2	3395	83.3	2934	8 AK120560	AK120560 Oryza sat
3	3155	77.4	2665	8 AY165599	AY165599 Saccharum
4	3097	76.0	2378	8 AK099716	AK099716 Oryza sat

5	3052.5	74.9	136267	8	AP005756	AP005756 Oryza sat
6	2825.5	69.3	2800	8	AK102640	AK102640 Oryza sat
7	2808	68.9	2614	8	HVU534445	AJ534445 Hordeum v
8	2674	65.6	2601	6	AR208568	AR208568 Sequence
9	2657	65.2	2570	8	AY094465	AY094465 Arabidops
10	2584	63.4	2190	6	AX506620	AX506620 Sequence
11	2584	63.4	2190	8	ATH532570	AJ532570 Arabidops
12	2583	63.4	2368	8	AK065191	AK065191 Oryza sat
13	2514	61.7	142114	8	AC073166	AC073166 Oryza sat
14	2514	61.7	300957	8	AE017116	AE017116 Oryza sat
15	2502	61.4	86710	8	ATF23E12	AL022604 Arabidops
16	2502	61.4	197859	8	ATCHRIV83	AL161587 Arabidops
17	2368	58.1	135583	8	AC136843	AC136843 Oryza sat
18	2275	55.8	2205	6	AX412656	AX412656 Sequence
19	2275	55.8	2205	6	AX507559	AX507559 Sequence
20	2245	55.1	2190	8	ATH532571	AJ532571 Arabidops
21	2217	54.4	2426	8	ATUGTRPR	Z50752 A.thaliana
22	2191.5	53.8	2516	8	HVU534446	AJ534446 Hordeum v
23	2144	52.6	103192	8	AC007369	AC007369 Arabidops
24	2144	52.6	143879	8	AC069251	AC069251 Genomic s
25	2128	52.2	114918	8	AP004945	AP004945 Lotus cor
26	2117	52.0	94349	8	ATF26013	AL133452 Arabidops
27	2093.5	51.4	100900	2	AC121239	AC121239 Medicago
28	1968	48.3	120743	8	AP004082	AP004082 Oryza sat
29	1944.5	47.7	115134	8	AC126786	AC126786 Medicago
30	1918	47.1	121112	8	AC144482	AC144482 Medicago
31	1670.5	41.0	1692	6	AR208569	AR208569 Sequence
32	1557.5	38.2	87592	8	AC099739	AC099739 Oryza sat
33	1557.5	38.2	127098	8	AC144426	AC144426 Oryza sat
34	1557.5	38.2	154128	8	AP000615	AP000615 Oryza sat
35	1470.5	36.1	1564	8	AK067391	AK067391 Oryza sat
36	1436	35.2	2661	8	AK120270	AK120270 Oryza sat
37	1424	34.9	1487	6	AR208571	AR208571 Sequence
38	1411	34.6	12809	8	AC127019	AC127019 Medicago
39	1409	34.6	2516	8	AK068224	AK068224 Oryza sat
40	1406	34.5	1625	8	AK059705	AK059705 Oryza sat
41	1174	28.8	1315	8	BT009593	BT009593 Triticum
42	1159.5	28.5	70311	8	AF128457	AF128457 Oryza sat
43	1140	28.0	77605	8	AF119222	AF119222 Oryza sat
44	1140	28.0	142852	8	AF161269	AF161269 Oryza sat
45	973	23.9	106246	8	AF528565	AF528565 Zea mays

ALIGNMENTS

RESULT 1
AR208565
LOCUS AR208565 2824 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 1 from patent US 6383776.
ACCESSION AR208565
VERSION AR208565.1 GI:21509752
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2824)
AUTHORS Allen,S.M., Hitz,W.D., Kinney,A.J. and Tingey,S.V.
TITLE Plant sugar transport proteins
JOURNAL Patent: US 6383776-A 1 07-MAY-2002;
FEATURES Location/Qualifiers
source 1..2824
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Alignment Scores:
Pred. No.: 3.68e-225 Length: 2824
Score: 3517.00 Matches: 684
Percent Similarity: 95.18% Conservative: 27
Best Local Similarity: 91.57% Mismatches: 34
Query Match: 86.31% Indels: 2
DB: 6 Gaps: 2

US-10-051-909-32 (1-800) x AR208565 (1-2824)

QY 56 MetSerGlyAlaValLeuValAlaValAlaSerIleGlyAsnLeuLeuGlnGlyTrp 75
 Db 238 ATGGGGGCGCGCTGATGGTCCCATCGCGGCTCTATCGCAACTTGTCTGAGGGCTGG 297
 QY 76 AspAsnAlaThrIleAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsn 95
 Db 298 GACATGCGACAATTGCTGGAGCGCTCTGTATACATAAAGAAGAAATTCACCTGCAGAGC 357
 QY 96 GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115
 Db 358 GAGCCTCTGATCGAAGGCTCATCTGCGCCATGTTCTCTCATTTGGGGCAACATCATACA 417
 QY 116 ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer 135
 Db 418 ACATCTCCGGGCGCAAGGCTGACTGCGTGGTAGGAGGCCCATGCTGGTCCCTCGGCT 477
 QY 136 IleLeuTyrPhePheSerGlyLeuIleMetLeuTyrSerProAsnValTyrValLeuLeu 155
 Db 478 GTCTCTACTTCGTGAGGCTGGTGTATCGGTTGGCGGCAATTGTGTACATCTTGCTC 537
 QY 156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr 175
 Db 538 CTCGAAAGGCTCATTTGATGGGTTGGTATCGGTTGGCGGTACACACTTGTCTCTCTAC 597
 QY 176 IleSerGluIleAlaProSerGluIleArg---GlyLeuLeuAsnThrLeuProGlnPhe 194
 Db 598 ATCTCCGAAACTGCACCGCACAGANATTCTTGGGCTGNTNGAACACGTTGCCGAGTTC 657
 QY 195 SerGly---SerGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSer 213
 Db 658 ATTGGGTCAGNGAGGATGTTCTCTCTCTACTGCGATGGTGTGGGATGTCCTCATG 717
 QY 214 ProSerProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePhe 233
 Db 718 CCCAAACCTGATTGGAGGCTCATGTTGGAGTTCTGTGATCCCGTCACTTATNTACTTT 777
 QY 234 GlyLeuThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMet 253
 Db 778 GGAATGACTGTCTTCTACTGCTGATCAACCAAGGTGGCTTGTNAGCAAAGGAGATG 837
 QY 254 AlaGluAlaLysLysValLeuGlnLysLeuArgGlyLysAspValSerGlyGluLeu 273
 Db 838 GCGAGCGGAAGAGAGTGTCAAGGCTCGCGGAAGAGAGATGTCTCANGGAGANG 897
 QY 274 SerLeuLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIle 293
 Db 898 GCTCTTCTAGTTGAAGGTTTGGGGTTCGGTAAAGATACACGATTTTNNAGATACATCAT 957
 QY 294 GlyProAlaThrGluAlaAlaAspLeuValThrAspGlyAspLysGluGlnIleThr 313
 Db 958 GGACCTGCCACCGAGGAGCCGATGATCTTGTAACTGACGGTGTATAGGAACAATCACA 1017
 QY 314 LeuTyrGlyProGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMet 333
 Db 1018 CTTTATGGGCTGAAGAGGCCAGTCAATGGATGCTCGACCTTCTAAGGGACCCATCATG 1077
 QY 334 LeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValPro 353
 Db 1078 CTTGGAAGTGTGCTTCTCTGATCTCGTCATGGGAGCATGGTGAACAGAGTGTACCC 1137
 QY 354 LeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly 373
 Db 1138 CTTATGGATCCGATTGTGACACTTTTGGTAGTGTCCATGAGAAATATGCTCAAGCTGGA 1197
 QY 374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393
 Db 1198 GGAAGTATGAGGAGCACATTTTCCAACTTTGGAAGTATGTTTCAGTGTACAGATCAG 1257
 QY 394 HisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluTyrAla 413
 Db 1258 CATGCCAAAATGACAGTGGGATGAAGAGAAATCTTCATAGGGATGACGAGGAGTACGCA 1317

QY 414 SerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGln 433
 Db 1318 TCTGATGGTGCAGGAGGTACTATGAGGACAAATCTCCATAGCCCATTTGCTGTCCAGGCAG 1377
 QY 434 AlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeuSer 453
 Db 1378 GCAACAGGTGCGGAAGGAAGGACATTTGTGCACCATGGTCAACCGTGAAGTCTTTGAGC 1437
 QY 454 MetArgArgGlnThrLeuLeuGlyGluGlyGlyAspGlyValSerSerThrAspIleGly 473
 Db 1438 ATGAGAAAGGCAAGCCCTCTTAGGGGAGGTGGAGATGGTGTGAGCAGCACTATATCGGT 1497
 QY 474 GlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArgLysGlu 493
 Db 1498 GGGGATGGCAGCTTGTCTGGAATGGTTCAGAGAAGGAAGGTGAGAAATGGTAGAAAGGAA 1557
 QY 494 GlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgGlySer 513
 Db 1558 GGTGGTTTCAAAAGAGTCTACTTGCACCAAGAGGAGTTCCTGGCTCAAGAGGGGCTCA 1617
 QY 514 IleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaAla 533
 Db 1618 ATTGTTTCACTTCCCGTGGTGGCGATGTTCTTGAAGGTAGTGTGAGTTGTACATGCTGCT 1677
 QY 534 AlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAsp 553
 Db 1678 GCTTTAGTAAGTCAGTCAGCACTTTTCTCAAGSGTCTTGTCTGAACCCATGTCAGAT 1737
 QY 554 AlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuPhe 573
 Db 1738 GCTGCCATGGTTCACCCATCTGAGGTAGCTGCAAGGTTTCAAGTGGAAAGATTTGTTT 1797
 QY 574 GluProGlyValArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnGlnPhe 593
 Db 1798 GAACCTGGAGTGAGCGTGCCTGTTAGTCCGTTGGTGAATTCAGATCCTTCAACAGTTT 1857
 QY 594 AlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyValAla 613
 Db 1858 GCTGGAATAAACGGTGTCTGTACTATATACCCCAAAATTTCTTGAGCAAGCTGGTGGCA 1917
 QY 614 ValIleLeuSerLysPheGlyLeuSerSerAlaSerIleLeuIleSerSerLeu 633
 Db 1918 GTTATTCTTTCCAAATTTGGTCTCAGCTCGCATCAGCATCCATCTTGTATCAGTCTCTC 1977
 QY 634 ThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGly 653
 Db 1978 ACTACCTTACTAATGCTTCTTGGCATTTGGCTTGGCATGCTGTATGATCTCTAGTATCT 2037
 QY 654 ArgArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuVal 673
 Db 2038 AGAAGGTTTTTGTCTGCTAGGCACAAATTCCAATCTTGTATAGCATCTCTAGTATCTCT 2097
 QY 674 ValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerVal 693
 Db 2098 GTGTCCAATCTAATGATTTGGGTACACTAGCCCATGCTTTGTCTCTCCACCATCAGTGT 2157
 QY 694 IleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAla 713
 Db 2158 ATCGTCTACTTCTGCTGCTTCTGTTATGGGATTTGGTCCCATGCTTTGTCTCTCCACCAT 2217
 QY 714 GluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrp 733
 Db 2218 GAGATCTTTCCAAACAGGCTTCGTGGCTCTGTATTTGCCATTTGTGCTTTACATTTCTGG 2277
 QY 734 IleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAla 753
 Db 2278 ATCGGAGATATCATCGTCACCTACAGCCTTCTGTGTGTGATGTGAATGCTATTTGGAGT 2337
 QY 754 GlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeuLys 773
 Db 2338 GGTGTTTTTCAGCATATATGAGTCGTATGCTTGTATTTCTTTGTTGTTGCTTCTTCTTAAG 2397

[illegible]

COMMENT

URL : <http://cdna01.dna.affrc.go.jp/cDNA/>

NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., Yamamoto, M. and Nakahama, Y.

FAIS Genome Sequencing & Analysis Group: Otomo, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiroaka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Takami-Takeda, Y., Tagawa, A., Takahashi, F., Takagi-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.

ORIGIN

```

Alignment Scores:
Pred. No.:      5.45e-217      Length:      2934
Score:          3395.00        Matches:     676
Percent Similarity: 86.84%    Conservative: 43
Best Local Similarity: 81.64% Mismatches:     76
Query Match:     83.31%      Indels:      34
DB:              8           Gaps:         5

US-10-051-909-32 (1-800) x AK120560 (1-2934)

Qy      3  SerGlySerTrpLeuAlaValGlnThrProPhethrProAspLeuAspArgArgGluArg 22
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      38 AGTCGTGGTGGCTCTTGCGCGACGCCGC---GCCCGCGCACCGTGCGCGCGTGCGCGGT 94
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      23 LeuLeuProSerVal-----ValLeuAla 30
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      95 GCGCTGTGCTACCGCCACGAGAGAGATCCGATCGCCCGCGCGCGCGGTGTTCTCTGCT 154
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      31 Leu-----ProGlyProLeuProProAlaSerCysSerSerGlnGluPro----- 45
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      155 CTTCCTCCTCCGCTGGGCCGACAAAGCGGGATCTGCGGGCGCCTT-CCTCCAGGGGAT 213
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      46 -----ValThrSerAsp-----Asp 50
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      214 CTCGCTCGCGCGCGCGCCCTCGCTCGTCTCGTCCGCGCAAGAGGTGGAAGAGGGGAT 273
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      51 IleLeuGluAspLysMetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsn 70
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      274 CGGAGAGGGGTGAAGATGTCGGGTGCTGCTCTCGTTCGGAATCGCGGCCCTCCATCGGCCAAC 333
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      71 LeuLeuGlnGlyTrpAspAsnAlaThrIleAlaAlaValLeuTyrIleLysLysGlu 90
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Db 334 CTGCTGAGGGATGGGACAAATGACCACTTGCAGGTGCTGTACTATACATAAAGGAA 393
 Qy 91 PheGlnLeuGlnAsnGluProThrValGluGlyLeuValSerMetSerLeuIleGly 110
 Db 394 TTCAAGCTAGAAAGTAGCCCACTGTGGAGGGGCTAATCGTGGCCATGTCACTGATTGGT 453
 Qy 111 AlaThrIleValThrThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMet 130
 Db 454 GCAACCATCATCAACAATCTCAGGGCCGGTATCAGACTGGATCGGCCGGCCCTATG 513
 Qy 131 LeuIleLeuSerSerIleLeuTyrPhePheSerGlyLeuIleMetLeuTyrSerProAsn 150
 Db 514 CTCATTCTCTTCAATTCTCTACTTCTCAGCAGCCTCATCATGCTATGGTCCCCGAAT 573
 Qy 151 ValTyrValLeuLeuLeuAlaArgPheValAlaAspGlyPheGlyIleGlyLeuAlaValThr 170
 Db 574 GTCTATGTGTACTGCTCGCAGCCCTTATAGATGGATCGGCATCGGCTTGGCTGTCA 633
 Qy 171 LeuValProLeuTyrIleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThr 190
 Db 634 CTTGTACCTTTGTACATCTCAGAGACAGCTCCTTCAGAGATCAGGGGTTTGTGAATACA 693
 Qy 191 LeuProGlnPheSerGlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMet 210
 Db 694 CTGCCACAGTTTCACTGGGTGAGGAGGATGTTCTTGTCTACTGCTGCTGCTGGGATG 753
 Qy 211 SerLeuSerProSerProAspTyrArgIleMetLeuGlyValLeuAlaIleProSerLeu 230
 Db 754 TCACTGTTGCCGTCACTGACTGGAGAAATATGCTTGGTGTCTCGCAATCCCTTCGTTG 813
 Qy 231 PhePhePheGlyLeuThrIlePheTyrLeuProGluSerProArgTyrPheValSerLys 250
 Db 814 TTTTCTTCGGATTGACAAATATCTATCTGCCAGAAATCACCAAGATGGTGTGTCAGCAA 873
 Qy 251 GlyArgMetAlaGluAlaLysLysValLeuGlnLysLeuArgGlyLysAspValSer 270
 Db 874 GGCGGATGGCTGAGGCAAGAGGTATTACAAAATACGTGGGAGAGAGATGTCTCA 933
 Qy 271 GlyGluLeuSerLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGlu 290
 Db 934 GGAGAAATGGCTCTCTTGTGAAGGTTTGGAGGTGGAGCTGACACTTCCATTGAAGAG 993
 Qy 291 TyrIleIleGlyProAlaThrGluAlaAlaAspLeuValThrAspGlyAspLysGlu 310
 Db 994 TACATCATCGACCGGCTATAGAGCCAGCTGATGAGCATGTTGTTGATGCGGATAGGAC 1053
 Qy 311 GlnIleThrLeuTyrGlyProGluGlyGlnSerTyrPheIleAlaArgProSerLysGly 330
 Db 1054 CAAATAACACTGTACGGGCTGTAGAGGCGCAATCATGATGCTCGACCTTCCAAGGA 1113
 Qy 331 ProIleMetLeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGln 350
 Db 1114 CCCAGCATCTTGGAGGTGCTTCTCTTACATCTCGTCAATGCGCAGCATGGTGAACCA 1173
 Qy 351 SerValProLeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetPro 370
 Db 1174 AGCGTCCACTTATGGATCCTATAGTCACGCTTTTCGGCAGTGTCCATGAGAACATG 1233
 Qy 371 GlnAlaGlyGlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerVal 390
 Db 1234 CATGCTGGAGGAAGCATGGGAGTACATTTGTTTCCCACTTTCGCTAGTATGTTAGTG 1293
 Qy 391 ThrAspGlnHisAlaLysAsnGluGlnTyrAspGluGluAsnLeuHisArgAspGlu 410
 Db 1294 ACAGATCAGACCCCCAAGGTTGATCAATGGGATGAGGAGAACCTTTCATAGGGATGATG 1353
 Qy 411 GluTyrAlaSerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeu 430
 Db 1354 GAATACGGCTGTGATGCTGGAGGTGACTATGAAGACAAATGTCCACAGCCCGTTGCTG 1413
 Qy 431 SerArgGlnAlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySer 450

Db 1414 TCCCGACAGACCACGAGCGCAGAAAGGAGGACATTCACACCATGCTCATCTCGGAAGT 1473
 Qy 451 AlaLeuSerMetArgArgGlnThrLeuLeuGlyGlyGlyGlyGlyValSerSerThr 470
 Db 1474 GCTCTGAGTATGAGAGAGAGAGCCCTCTTGGAAAGAGGGTGGCGGGGTGAGCAGCACT 1533
 Qy 471 AspIleGlyGlyGlyTyrPheGlnLeuAlaTyrLysTyrSerGluLysGluGlyGluAsnGly 490
 Db 1534 GGCATTGGTGGGGATGGCAGCTTGCATGGAAATGGTCAGAGCGAGAGGCGGAGATGGT 1593
 Qy 491 ArgLysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArg 510
 Db 1594 AAGAAGGAAGGAGGATTTAAAGAGGATCTACTTGCACCAAGAGGAAGTCCAGCATCAAGA 1653
 Qy 511 ArgGlySerIleValSerLeuProGlyGlyGlyGlyGlyValPheGluGlySerGluPheVal 530
 Db 1654 AGGGGCTCAGTTATTTCACTTCTGTTGGAGGGAGCGCTCCTGAAGGCGAGCAATTCATA 1713
 Qy 531 HisAlaAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArg 550
 Db 1714 CATGCTGCCGCTTTGGTAAGCCCAACAGCAGCTTTACTCCAAAGGATATCATTTGAACAGCGT 1773
 Qy 551 MetSerAspAlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTyrLys 570
 Db 1774 ATGTCGGTCCAGCCATGATTCATCCATCAGAGGAGCTGCCAAAGGTTCAAGCTGGA 1833
 Qy 571 AspLeuPheGluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeu 590
 Db 1834 GATTGTTTGAACCTGGAGTGAGCGGTGCTGTTGTAGTGGTGTGGAAATTCAAATCCTT 1893
 Qy 591 GlnGlnPheAlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGlnAla 610
 Db 1894 CAACAGTTTGTGGAATAAATGGGTTCTCTATTACACTCCACAAATACTTGAGCAAGCG 1953
 Qy 611 GlyValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIle 630
 Db 1954 GGGGTGGCAGTTCTCCTTTCCAATCTTGGCCCTCAGTTCAGCATCAGCTTCCATCTTGATC 2013
 Qy 631 SerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAsp 650
 Db 2014 AGTCTCTGACCCCTACTGATGCTTCTAGCATTTGGTTTAGCCATGAGACTTATGGAC 2073
 Qy 651 LeuSerGlyArgArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuVal 670
 Db 2074 ATCTCTGGAAGAGGTTTCTGCTTCTGGGCACAAATCCAGTCTTGTATAGCATCTCTAGTT 2133
 Qy 671 IleLeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThr 690
 Db 2134 GTCTTGGTGTGTCCAAATGTTATCGACCTGGGTACAGTGGCCCAACGCCGACCTCCACA 2193
 Qy 691 ValSerValIleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIle 710
 Db 2194 ATCAGCGTCATCATCTACTTCTGCTGCTTCTGTCATGGGATTCGGTCCGATCCCAACATT 2253
 Qy 711 LeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPhe 730
 Db 2254 CTGTGTGCGGAGATCTTCCCACTAGGGTCCGCGGCATCTGCATTGCCATCTCGCCCTG 2313
 Qy 731 ThrPheTyrIleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIle 750
 Db 2314 ACATTCTGGATTGGTATATCATAGTACACCTACAGCCTCCCTGTGATGCTGAATGCCATC 2373
 Qy 751 GlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheVal 770
 Db 2374 GGCCTAGCAGGTGCTTGTGATATATACCGAGTGTGCTCGATTGCTTCTGTTGCTGTC 2433
 Qy 771 PheLeuLysValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAla 790
 Db 2434 TTCCTCAAGGTCCCGAGACGAGGCGATGCGGCTCGAAGTCAATCACCAGGATTTCTTCGCT 2493
 Qy 791 ValGlyAlaLysGlnAlaAla 798
 Db 2494 GTTGGTGCAAGCAAAATGCAGGCT 2517

RESULT 3
AY165599
LOCUS
DEFINITION
Saccharum hybrid cultivar putative sugar transporter type 2a mRNA, complete cds.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Saccharum hybrid cultivar
Saccharum hybrid cultivar
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum complex.
REFERENCE
1 (bases 1 to 2665)
Casu,R.E., Grof,C.P.L., Rae,A.L., McIntyre,C.L., Dimmock,C.M. and Manners,J.M.
Identification of a novel sugar transporter homologue strongly expressed in maturing stem vascular tissues of sugarcane by expressed sequence tag and microarray analysis
Plant Mol. Biol. 52, 371-386 (2003)
REFERENCE
2 (bases 1 to 2665)
Casu,R.E., Grof,C.P.L., Rae,A.L., McIntyre,C.L., Dimmock,C.M. and Manners,J.M.
Direct Submission
Submitted (17-OCT-2002) Division of Plant Industry, Commonwealth Scientific and Industrial Organisation, 120 Meiers Rd, Indooroopilly, QLD 4068, Australia
FEATURES
source
location/Qualifiers
1..2665
/organism="Saccharum hybrid cultivar"
/mol_type="mRNA"
/db_xref="taxon:128810"
97..2334
/codon_start=1
/product="putative sugar transporter type 2a"
/protein_id="AA037640.1"
/db_xref="GI:31505504"
/translation="MSGAAIVAIIVASIGNLLQWDNATISAALLYIKKEFKLESEPTV EGLIVAMSLIGATIIITFSGPVSDWIGRLPMLILSSVLYFVSSLIMLWSPNYYVLLLA RLINGFGVGLAVTLVPLYISETAPPEIRGLLNTLPQPSGGMFLSYCMVFGMSLLPS PDWRIMGLVALPSPFFFGUTIFYLPESPRWLVSGRMAEAKVLQRLRGKDDVTGEM ALLVBEGLVGDDTFIEEYITGPAIEPADDHVADGDRERIILFGPEEQSWIARPSKGS SMLGSVLSIGSRHGSVMNQNLPLTDPIVTLFGSVHENISLAGGSMRRIVPSKFDMSIS ITDQHPKTEQWDEENLHRDNEEYATDGAASDYENVHSPLLSRQNTSAEGKIAHHGH RGSALGLRRRLSDEGGEATSTGIGGGWQLAWKSEREGEDGKKEGSFKRIYLVHQEG VAGSRGVSVSLPGGSDASEGKFIHAAALVSQSALYPRDITEQRMAGPAMWHPSEAA AKVPSWRDLFEPGVRALLVIGIGIQLQQFAGINGVLYYTPQIMEQAGVAVLISNLGL SSASASILSSVTALLMPLSGLAMRLMDVSGRRFLLSTIPVLIASLIVLVSNVIE LGTVHVALVSTISVITYLCCFKMGFGPIPNILCAEFPTVRVGICIAICALIFWVGDI IVTVSLPVMNLNAIGLEGVFGIYAVACAIASFVFLVKVPETKGMPLVITEFFAVGAKQ AVAKA"
97..156
/note="putative; transmembrane-region site"
115..774
/note="pfam00083; Region: sugar transporter"
238..297
/note="putative; transmembrane-region site"
313..372
/note="putative; transmembrane-region site"
406..462
/note="putative; transmembrane-region site"
511..570
/note="putative; transmembrane-region site"
589..654
/note="putative; transmembrane-region site"
658..1794
/note="Region: large cytoplasmic loop"
1624..2298
/note="pfam00083; Region: sugar transporter"
1672..1731
/note="putative; transmembrane-region site"

misc_feature 1813..1869
/note="putative; transmembrane-region site"
misc_feature 1897..1953
/note="putative; transmembrane-region site"
misc_feature 1969..2022
/note="putative; transmembrane-region site"
misc_feature 2089..2151
/note="putative; transmembrane-region site"
misc_feature 2191..2247
/note="putative; transmembrane-region site"
ORIGIN
Alignment Scores:
Pred. No.: 5.02e-201 Length: 2665
Score: 3155.00 Matches: 601
Percent Similarity: 89.56% Conservative: 68
Best Local Similarity: 80.46% Mismatches: 78
Query Match: 77.42% Indels: 0
DB: 8 Gaps: 0
US-10-051-909-32 (1-800) x AY165599 (1-2665)
QY 54 AsPLysMetSerGlyAlaValLeuValAlaLeValAlaSerIleGlyAsnLeuLeuGln 73
Db 91 GAAACATGTCGGAGCTGCTCTTTGTTGCAATAGTAGCTTCGATCGCAATCTGCTTCAG 150
QY 74 GlyTrpAspAsnAlaThrIleAlaAlaValLeuTyrIleLysLysGluPheGlnLeu 93
Db 151 GGGTGGGACATGCCACCATCTCAGCTGCTCTATTATATATAAAGAGGAATTCAAACTA 210
QY 94 GlnAsnGluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIle 113
Db 211 GAAAGTGAGCCTACTGTGGAGGGCTGATTGTGGCCATGTCAATAATGGAGCAACTATC 270
QY 114 ValThrThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeu 133
Db 271 ATTACAACATTTCTGTGACCATATCAGACTGATGGCCGCTCCCTATGCTCATCCTC 330
QY 134 SerSerIleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyrVal 153
Db 331 TCTTCAGTTCTGFACTTTGTTAGCAGCCTCATTTATGCTATGTTGGTCCCGAATGTTATGTA 390
QY 154 LeuLeuLeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValPro 173
Db 391 CTACTGTGGCACGCTTATAAATGGATTTCGTTGCGTTCGGCTTGGCTGTCACCTTTGTTTCT 450
QY 174 LeuTyrIleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGln 193
Db 451 CTGTACATTTTCAGAAACAGCTCCTCCAGAGATAAGAGTTTGTCTGAATACACTGCCACAG 510
QY 194 PheSerGlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSer 213
Db 511 TTCAGTGGGTACGGGGGATGTTTTTGTCTACTGCATGGTGTTCGGGATGTCACCTGCTG 570
QY 214 ProSerProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePhePhe 233
Db 571 CCATCACCTGATTGGAGGATTATGCTTGGGGTTCTCGCGTTACCTTTCATTTGTTTCTTT 630
QY 234 GlyLeuThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMet 253
Db 631 GGTTTGACAAATATTTTATCTGCCTGAATCACCAAGATGGCTTGTTCAGCAAGGCGGAATG 690
QY 254 AlaGluAlaLysLysValLeuGlnLysLeuArgGlyLysAspValSerGlyGluLeu 273
Db 691 GCAGAGGCAAGAGGTGTGCAAGATTACGTGGGAAGGATGACGTCACAGCGCGAAATG 750
QY 274 SerLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIle 293
Db 751 GCTCTTCTTGTGAAGGGTTGGAGGTTCGGAGGGGATACCTTTCATTGAAGAGTACATCATT 810
QY 294 GlyProAlaThrGluAlaAlaAspLeuValThrAspGlyAspLysGluGlnIleThr 313
Db 811 GGTCTGCTATCGAACCCAGCTGATGACCACGTCGCTGATGGTGTATAGAACGAATAATA 870


```
QY 314 LeuTyrGlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMet 333
|||:::|||||
Db 871 CTATTGGGCTGAAGAGGGCCAAATCATGGATAGCTCGGCTTCCAAAGGGTCCAGCATG 930

QY 334 LeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValPro 353
|||:::|||||
Db 931 CTGGAAAGTGTGCTTTCTATCGGATCTCGTCATGGCAGCATGGTGAACCAAAATTTACCC 990

QY 354 LeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly 373
|||:::|||||
Db 991 CTTACGGATCCTATTGTACAGCTTTTGGAGTGTCCATGAGAACATCTCTCTCGCTGGA 1050

QY 374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393
|||:::|||||
Db 1051 GGAAGTATGAGGAGAAATAGTGCTTCTAAATTTGACAGTATGATCAGTATCACAGATCAG 1110

QY 394 HisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAla 413
|||:::|||||
Db 1111 CACCCCAAGACCGAGCAGTGGGATGAGGAGAACCTCCATAGGGATAATGAGGAGTATGCA 1170

QY 414 SerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGln 433
:::|||||
Db 1171 ACTGATGGTGTGCAAGTGAAGTACTATGAAACAATGATGACACAGCCCTCTGCTCTCCGACAG 1230

QY 434 AlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeuSer 453
|||:::|||||
Db 1231 AATACAAGTCCCGAAGGAGGGCATTTGCACATCAGGGCACCTGTGAAGTCTTTGGGT 1290

QY 454 MetArgArgGlnThrLeuLeuGlyGluGlyGlyAspGlyValSerSerThrAspIleGly 473
:::|||||
Db 1291 TTGAGAAGAAGAGCCTGTCCGATGAGGTGGTGGTGGCAGCAACCAGCAGCTGGCATTGGT 1350

QY 474 GlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArgLysGlu 493
|||:::|||||
Db 1351 GGCGGATGCCAGCTCGCCTGGAAATGGTCAGAGCGGAGAGGTGAGGACGGTAAGAAGGAA 1410

QY 494 GlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArgGlySer 513
|||:::|||||
Db 1411 GGGAGTTTCAAAGAATCTACTTACACCAAGAGGAGTGTGCTGGCTCAAGAAGGGGATCT 1470

QY 514 IleValSerLeuProGlyGlyAspValPheGluGlySerGluPheValHisAlaAla 533
:::|||||
Db 1471 GTTGTCTCACTTCTCTGGTGGAGGTGATGCCTCTGAAGGTGGTAAGTTTATACATGCTGCT 1530

QY 534 AlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAsp 553
|||:::|||||
Db 1531 GCTCTGGTCAGCCAGTCAGTCTCTTACCCGAGGGATATTACAGAACAGCCGATGGCCGGT 1590

QY 554 AlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuphe 573
|||:::|||||
Db 1591 CCAGCTACAATGCACCCATCAGAGGCAGCTGCCAAAGTTTCCAAAGCTGGAGAGACCTGTTT 1650

QY 574 GluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnPhe 593
|||:::|||||
Db 1651 GAACCTGGTGTGAGACGTGCCCTGTTGTTGGTGGATTTGAATTCAGATCCTTCAACAGTTC 1710

QY 594 AlaGlyIleAsnGlyValLeuTyrTrpThrProGlnIleLeuGluGlnAlaGlyValAla 613
|||:::|||||
Db 1711 GCGGGAATAAATGGGTTCTCTACTACACTCCGAAAATTATGGAGCAAGCTGGTGTGGCA 1770

QY 614 ValIleLeuSerLysPheGlyLeuSerSerAlaSerIleLeuIleSerSerLeu 633
|||:::|||||
Db 1771 GTTCTTATTCCAAATCTTGGTCTCAGTTCGGCATCAGCATCCATCCTTAATAGTTCTGTC 1830

QY 634 ThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGly 653
|||:::|||||
Db 1831 ACTGCCCTACTTATGCTCCCTAGCATTTGGTTAGCCATGAGACTTATGGACGTATCTGGA 1890

QY 654 ArgArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuVal 673
|||:::|||||
Db 1891 AGAAGGTTTCTGCTGCTAAGCACAAATTCACAGTCTTGATAGCATCTTAAATGTCCTGGTT 1950
```

```
QY 674 ValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerVal 693
|||:::|||||
Db 1951 GTGTCTAATGTTATCGAGTGGGTACAGTGGTCCATGCTGTGCTCTCCACAATTAGTGTG 2010

QY 694 IleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAla 713
|||:::|||||
Db 2011 ATCACTACTTATGCTGCTTCAAAATGGCTTTGGCCCCATTTCCCAACATTTCTATGTGCA 2070

QY 714 GluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrp 733
|||:::|||||
Db 2071 GAGTTTTTCCAAACCAGGGTTCGTGGCATCTGCAATGCCATCTGTGCCTTGATATTTGG 2130

QY 734 IleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAla 753
:::|||||
Db 2131 GTTCGAGACATTATTGTGCATACATACAGTCTTCTCTGTGATGCTGAATGCTATTGGTCTAGAA 2190

QY 754 GlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeuLys 773
|||:::|||||
Db 2191 GCGGTGTTGGCATATACGAGTTGCTGTGCCATTCCTTTGTTGTTGTTGTTGTTGTTAAG 2250

QY 774 ValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAla 793
|||:::|||||
Db 2251 GTTCCGGAGACAAAGGGCATGCCCTTGAGGTCTACTGAGTCTTTTGCAGTTGGTGCA 2310

QY 794 LysGlnAlaAlaAlaLysAla 800
|||:::|||||
Db 2311 AAGCAAGCGGTTCGAAAGGCC 2331

RESULT 4
AK099716 2378 bp mRNA linear PLN 24-JUL-2003
LOCUS Oryza sativa (japonica cultivar-group) cDNA clone:J013086B19, full
DEFINITION insert sequence.
ACCESSION AK099716
VERSION AK099716.1 GI:32984925
KEYWORDS FLJ_CDNA; CAP trapper.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1
AUTHORS The Rice Full-Length cDNA Consortium, National Institute of
Agrobiological Sciences Rice Full-length cDNA Project Team:,
Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K.,
Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I.,
Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C.,
Ohtsuki,K., Shishiki,T., Foundation of Advancement of International
Science Genome Sequencing & Analysis Group:, Otomo,Y., Murakami,K.,
Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y.,
Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M.,
Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niihara,J.,
Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J.,
Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN:,
Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S.,
Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M.,
Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y.,
Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
Yoshino,M. and Hayashizaki,Y.
Collection, mapping, and annotation of over 28,000 cDNA clones from
japonica rice
Science 301 (5631), 376-379 (2003)
JOURNAL 22752273
MEDLINE 12869764
PUBMED 2 (bases 1 to 2378)
REFERENCE
AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K.,
Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T.,
Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K.,
Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I.,
Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M.,
Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M.,
Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
```


Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Namiki, T., Narikawa, R., Niikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Osato, N., Ota, Y., Otomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.
Direct Submission
Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)
This clone is one of the 28K full-length cDNA clones from japonica rice.
URL : http://cdna01.dna.affrc.go.jp/cDNA/
NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and Yamamoto, M.
FAIS Genome Sequencing & Analysis Group: Otomo, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.
Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.
Location/Qualifiers
1. .2378
/organism="Oryza sativa (japonica cultivar-group)"
/mol type="mRNA"
/cultivar="Nipponbare"
/db xref="taxon:39947"
/clone="J013086B19"

FEATURES
source

1. .2378
/organism="Oryza sativa (japonica cultivar-group)"
/mol type="mRNA"
/cultivar="Nipponbare"
/db xref="taxon:39947"
/clone="J013086B19"

ORIGIN

Alignment Scores:
Pred. No.: 3.17e-197 Length: 2378
Score: 3097.00 Matches: 596
Percent Similarity: 93.20% Conservative: 34
Best Local Similarity: 88.17% Mismatches: 46
Query Match: 76.00% Indels: 0
DB: 8 Gaps: 0

US-10-051-909-32 (1-800) x AK099716 (1-2378)

QY 123 AspSerIleGlyArcArgProMetLeuIleLeuSerSerIleLeuTyrPhePheSerGly 142
|||
Db 1 GACTGGATCGCGCGCCCTATGCTCAATTCCTCAATTCCTCAATTCCTCAGCAGC 60
QY 143 LeuIleMetLeuTrpSerProAsnValTyrValIleLeuLeuAlaArgPheValAspGly 162
|||
Db 61 CTCATCATGCTATGGTCCCGGAATGTCTATGTGTACTGCTCGCACGCCCTTATAGATGGA 120

QY 163 PheGlyIleGlyLeuAlaValThrLeuValProLeuTyrIleSerGluIleAlaProSer 182
|||
Db 121 TTCGGCATCGGCTTGGCTGTCACTTGTACCTTTGTACATCTCAGAGACAGCTCCTTCA 180
QY 183 GluIleArgGlyLeuLeuAsnThrLeuProGlnPheSerGlySerGlyMetPheLeu 202
|||
Db 181 GAGATCAGGGGTTTGGCTGAATACACTGCCACAGTTCAGTGGGTCAGGAGGGATGTTCTTG 240
QY 203 SerTyrCysMetValPheGlyMetSerLeuSerProSerProAspTrpArgIleMetLeu 222
|||
Db 241 TCCTACTGCATCGTGTGTTGGGATGTCACTGTTGCCGTCACTGCTGAGGAAATATGCTT 300
QY 223 GlyValLeuAlaIleProSerLeuPhePheGlyLeuThrIlePheTyrLeuProGlu 242
|||
Db 301 GGTGTTCTCGCAATCCCTTCGTTGTTTCTTCGGATTGACAAATATCTATCTGCCAGAA 360
QY 243 SerProArgTrpLeuValSerLysGlyArgMetAlaGluAlaLysLysValLeuGlnLys 262
|||
Db 361 TCACCAAGATGGCTTGTTCAGCAAGGGCGGATGGCTGAGGCAAGAGGTTATTACAAAA 420
QY 263 LeuArgGlyLysAspValSerGlyGluLeuSerLeuLeuGluLeuVal 282
|||
Db 421 TTACGTGGGAGAGAGGATGTCTCAGGAGAAATGGCTCTTCTTGTGTTGAAGGTTTGGAGGT 480
QY 283 GlyGlyAspThrSerIleGluGluTyrIleIleGlyProAlaThrGluAlaAlaAspAsp 302
|||
Db 481 GGAGCTGACACTTCCATTGAAGAGTACATCATCGGACCGGCTATAGAGCCAGCTGATGAG 540
QY 303 LeuValThrAspGlyAspLysGluGlnIleThrLeuTyrGlyProGluGluGlyGlnSer 322
|||
Db 541 CATGTTGTTGATGGCGATAAGGACCAATAACACTGTACGGGCTGAAGAGGGCCAATCA 600
QY 323 TrpIleAlaArgProSerLysGlyProIleMetLeuGlySerValLeuSerLeuAlaSer 342
|||
Db 601 TGGATTGCTCGACCTTCCAAGGGACCCAGCATTTCTTGAAGTGTGCTTCTTACATCT 660
QY 343 ArgHisGlySerMetValAsnGlnSerValProLeuMetAspProIleValThrLeuPhe 362
|||
Db 661 CGTCATGGCAGCATGGTGAACACGAGCGGTGCTTATGGATCCTATAGTACAGCTTTTC 720
QY 363 GlySerValHisGluAsnMetProGlnAlaGlySerMetArgSerThrLeuPhePro 382
|||
Db 721 GGCAGTGTCCATGAGAACATGCCTCATGCTGGAGGAAGCATGCGGAGTACATTTGTTCCC 780
QY 383 AsnPheGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTrpAspGlu 402
|||
Db 781 AACTTTGGTAGTATGTTTAGTGTGACAGATCAGCACCCCAAGTTGATCAATGGGATGAG 840
QY 403 GluAsnLeuHisArgAspAspGluGluTyrAlaSerAspGlyAlaGlyLysAspTyrGlu 422
|||
Db 841 GAGAACCTTCATAGGGATGATGAGGAATACGCTCTGATGGTGTGAGGTTGACTATGAA 900
QY 423 AspAsnLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIle 442
|||
Db 901 GACAATGTCCACAGCCGTTGCTGTCCCGACAGACCACGAGCGCAGAGGGAAGGACATT 960
QY 443 ValHisGlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeuGlyGlu 462
|||
Db 961 GCACACCATGCTCATCGTGAAGTGTCTCTGATATGAGAAGAAGCCCTCTTGAAGAG 1020
QY 463 GlyGlyAspGlyValSerSerThrAspIleGlyGlyGlyTrpGlnLeuAlaTrpLysTrp 482
|||
Db 1021 GGTGGCAGGCGGTGACGACACTGGCATTTGTGGGGATGGCAGCTTGATGGAATGG 1080
QY 483 SerGluLysGluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyrLeuHis 502
|||
Db 1081 TCAGAGCGAGAAGGCGGAGGATGGTAAGAAGGAAGGAGGATTAAAGGATCTACTTGCAC 1140
QY 503 GlnGluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGlyGlyAsp 522
|||
Db 1141 CAAGAGGAAGTCCAGGATCAAGAGGGGCTCAGTTATTTCATCTTCCTGGTGGAGGGGAC 1200

Db 51020 TGAGCATGTTGTTGATGGCGATAAGGACCAATAACACTGTACGGCCTGAAGAGGGCCA 51079
QY 321 nSerTrpIleAlaArgProSerLysGlyProIleMetLeuGlySerValLeuSerLeuAl 341
Db 51080 ATCATGGATTGCTCGACCTTCCAGGGACCCAGCAATCTTGGAAAGTGTGCTTCTCTTAC 51139
QY 341 aSerArgHisGlySerMetValAsnGlnSerValProLeuMetAspProIleValThrLe 361
Db 51140 ATCTCGTCATGGCAGCATGGTGAACAGAGCGTGCACCTTATGGATCCTATAGTCAGCCT 51199
QY 361 uPheGlySerValHisGluAsnMetProGlnAlaGlyGlySerMetArgSerThrLeuPh 381
Db 51200 TTTCCGAGTGTCCATGAGAACATGCCTCATGCTGGAGGAACCATCGGGAGTACATGTT 51259
QY 381 eProAsnPheGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTrpAs 401
Db 51260 TCCCAACTTTGGTAGTATGTTTAGTGTGACAGATCAGCACCCCAAGGTTGATCAATGGGA 51319
QY 401 pGluGluAsnLeuHisArgAspGluGluTyrAlaSerAspGlyAlaGlyGlyAspTy 421
Db 51320 TGAGGAGAACCTTCATAGGGATGATGAGGAATACGCGTCTGATGGTGTGGAGGTGACTA 51379
QY 421 rGluAspAsnLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAs 441
Db 51380 TGAAGACAATGTCCACAGCCCGTTGTGTCCCGACAGACCACGAGCGCAGAAAGGAAGGA 51439
QY 441 pIleValHisGlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeuG1 461
Db 51440 CATTCACACCATGCTCATCGTGGAGTGTCTGTAGTATGAGAAAGAAAGCCCTCTTGGA 51499
QY 461 yGluGlyGlyAspGlyValSerSerThrAspIleGlyGlyGlyTrpGlnLeuAlaTrpLy 481
Db 51500 AGAGGGTGGCGAGCGCGGTGAGCAGCACTGGCATTTGGTGGGGATGGCAGCTTGCGAA 51559
QY 481 sTrpSerGluLysGluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyrLe 501
Db 51560 ATGGTTCAGAGCGAGAAGCGGAGGATGTTAAGAAAGGAAGGAGGATTTAAAGGATCTACTT 51619
QY 501 uHisGlnGluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGlyG1 521
Db 51620 GCACCAAGAGGAAGTCCAGGATCAAGAGGGGCTCAGTTATTCTACTTCCTGGTGGAGG 51679
QY 521 yAspValPheGluGlySerGluPheValHisAlaAlaLeuValSerGlnSerAlaLe 541
Db 51680 GGACGCTCCTGAAGCGCAGCAATTTCATACATGCTGCCGCTTGGTAAGCCACCCAGCACT 51739
QY 541 uPheSerLysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisProSerG1 561
Db 51740 TTACTCCAAGGATATCATTTGAACAGCGTATGTCGGTCCAGCCCATGATTCATCCATCAGA 51799
QY 561 uValAlaAlaLysGlySerArgTrpLysAspLeuPheGluProGlyValArgArgAlaLe 581
Db 51800 GGCAGCTGCCAAAGGTTCAAGCTGGAAAGATTGTTTGAACCTGGAGTGAGCGGTGCGIT 51859
QY 581 uLeuValGlyValGlyIleGlnIleLeuGln----- 591
Db 51860 GTTAGTCGGTGTGGAAATTCAAATCCTTCAACAGGTAGAGATGATACCAATTCATAATT 51919
QY 591 ----- 591
Db 51920 CTTTTTTGTACTGTCCAAACATTGACGGTGTGTTAAGATTGTGTACTATTAAAGAAATTCACT 51979
QY 592 -----GlnPheAlaGlyIleAsnGlyValLeuTyrTyrT 603
Db 51980 AATTCTCTTTTATGTTCTTCTTGTGTCAGTTTGTCTGGAATAAAATGGGTTCTCTATTACA 52039
QY 603 hrProGlnIleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerS 623
Db 52040 CTCACAAATACTTGAGCAAGCGGGGTGGCAGTTCTCCTTTCCAAATCTTGGCCTCAGTT 52099
QY 623 erAlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIleG 643

Db 52100 CAGCATCAGCTTCCATCTTTGATCATGTTCTCTGTGACCACCCTACTGATGCTTCCCTAGCAATTG 52159
QY 643 lyPheAlaMetLeuLeuMetAspLeuSerGlyArg----- 654
Db 52160 GTTTAGCCCATGAGACTTATGGACATCTCTGGAAGAAGGTACTAACTTGTTCCTTGGTTT 52219
QY 654 ----- 654
Db 52220 GCATCATCTTAGTACTACCAAAATATACCAAAAGAAATGCACCCCAAGTTAGTATTAATT 52279
QY 655 -----ArgPheLeuLeuLeuGlyThrIleProIleLeuIleA 667
Db 52280 ACATCATACATTGTTTCCCTTTTAGGTTTCTGCTTCTGGGCACATTCAGTCTTGATAG 52339
QY 667 laSerLeuValIleLeuValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaL 687
Db 52340 CATCTCTAGTTGTTGTTGTGTCCAAATGTTATCGACCTGGGTACAGTGGCCACGCCG 52399
QY 687 euLeuSerThrValSerValIleValTyrPheCysCysPheValMetGlyPheGlyProI 707
Db 52400 CACTCTCCACAATCAGCGTCATCATCTACTTCTGCTTCTGTCATCGGATTCGGTCCGA 52459
QY 707 leProAsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaI 727
Db 52460 TCCCCAACATTCTGTGTGCGGAGATCTTCCCACTAGGGTCCGCGCATCTGCATTGCCA 52519
QY 727 leCysAlaPheThrPheTrpIleGlyAspIleIleValThrTyrSerLeuProValMetL 747
Db 52520 TCTGCGCCCTGACATCTCGATTGTTGGTATCATAGTACCTACAGCTCCCTGTGATGC 52579
QY 747 euAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeuIleSerP 767
Db 52580 TGAATGCCATCGGCTAGCAGGTGCTCTTGGTATATACGACAGTGTGTTGCTCGATTGCCT 52639
QY 767 heValPheValPheLeuLysValProGluThrLysGlyMetProLeuGluValIleThrG 787
Db 52640 TCGTGTTCGTCTTCTCAAGGTCCCCGAGACGAAGGGCATGCCGCTCGAAGTCATCACCG 52699
QY 787 luPhePheAlaValGlyAlaLysGlnAlaAlaAala 798
Db 52700 AGTTCTTCGCTGTTGTTGTCGCAAGCAAAATGCAGGCT 52734

RESULT 6
AK102640 2800 bp mRNA linear PLN 24-JUL-2003
LOCUS Oryza sativa (japonica cultivar-group) cDNA clone J033100A10, full
DEFINITION insert sequence.
ACCESSION AK102640
VERSION AK102640.1 GI:32987849
KEYWORDS FLI_CDNA; CAP trapper.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1
REFERENCE The Rice Full-Length cDNA Consortium, National Institute of
AUTHORS Agrobiological Sciences Rice Full-Length cDNA Project Team;
Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K.,
Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I.,
Koizima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C.,
Ohtsuki,K., Shishiki,T., Foundation of Advancement of International
Science Genome Sequencing & Analysis Group; Otomo,Y., Murakami,K.,
Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y.,
Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M.,
Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niihara,J.,
Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Maeda,J.,
Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN;
Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S.,
Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M.,
Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y.,
Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
Yoshino,M. and Hayashizaki,Y.

TITLE	Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice
JOURNAL	Science 301 (5631), 376-379 (2003)
MEDLINE	22752273
PUBMED	12869764
REFERENCE	2 (bases 1 to 2800)
AUTHORS	Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K., Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K., Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M., Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M., Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Kurosaki,T., Kusumegi,T., Li,C., Lu,M., Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A., Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakamura,M., Namiki,T., Narikawa,R., Niikura,J., Nishi,K., Nomura,K., Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H., Osato,N., Ota,Y., Otomo,Y., Ryu,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Shibata,K., Shishiki,T., Sogabe,Y., Sugano,S., Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,W., Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S. and Yoshimura,A.
TITLE	Direct Submission
JOURNAL	Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)
COMMENT	This clone is one of the 28K full-length cDNA clones from japonica rice. URL : http://cdna01.dna.affrc.go.jp/cDNA/ NIAS Rice Full-Length cDNA Project Team: Kikuchi,S., Sato,H., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T. and Yamamoto,M. FAIS Genome Sequencing & Analysis Group: Otomo,Y., Iida,Y., Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M., Kodama,T., Kurosaki,T., Kusumegi,T., Lu,M., Masuda,H., Miura,J., Mizuno,K., Narikawa,R., Niikura,J., Oka,M., Ryu,R., Sugano,S., Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S., Yoshimura,A., Matsubara,K. and Murakami,K. Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Iida,J., Imamura,K., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,J., Kishikawa-Hirozane,T., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Osato,N., Ota,Y., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Waki,K., Yasunishi,A. and Hayashizaki,Y. Location/Qualifiers 1. .2800 /organism="Oryza sativa (japonica cultivar-group)" /mol_type="mRNA" /cultivar="Nipponbare" /db_xref="taxon:39947" /clone="J033100A10"
FEATURES	source
ORIGIN	Alignment Scores: Pred. No.: 5.39e-179 Length: 2800 Score: 2825.50 Matches: 560

Percent Similarity:	82.32%	Conservative:	78
Best Local Similarity:	72.26%	Mismatches:	121
Query Match:	69.34%	Indels:	16
DB:	8	Gaps:	7
US-10-051-909-32 (1-800) x AK102640 (1-2800)			
QY	30	AlaLeuProGlyProLeuProPro-AlaSerCys-----SerSerG1	43
Db	117	GCCTTCCAGATTCCCGCGCTCTCTTCTTTAGGGGATCCGAAATCTCGGTGGACG	176
QY	43	nGluProValThrSerAspAspIleLeuGluAspLysMetSerGlyAlaValLeuValAl	63
Db	177	AGAGACTTGGTGGTAAGAT-----TCGCCGGCCATGGCGCGCTGCTGGTGC	227
QY	63	aIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrpAspAsnAlaThrIleAlaAl	83
Db	228	CATCGCGGCTCCATCGGCACTTGTCTGCAGGGCTGGGATATGCAACCATTCAGGTGC	287
QY	83	aValLeuTyrIleLysLysGluPheGlnLeuGlnAsnGluProThrValGluGlyLeuI	103
Db	288	GGTACTGTACATCAAGAAGGAATTCACCTTGCAGAGCGAGCCCTTATCGAAGSCCTGAT	347
QY	103	eValSerMetSerLeuIleGlyAlaThrIleValThrPheSerGlyProLeuSerAs	123
Db	348	CGTGGCCATGTCGCTCAITGGGGCGACGATCATCACCAGCTTCTCTGGAGCAGTGGCTGA	407
QY	123	pSerIleGlyArgArgProMetLeuIleLeuSerSerIleLeuTyrPhePheSerGlyLe	143
Db	408	TTCTTTTGGTAGGGGCCCATGCTGATCGCGTTCGGCTGTCTCTACTTTGTAGTGGCT	467
QY	143	uIleMetLeuTrpSerProAsnValTyrValLeuLeuLeuAlaArgPheValAspGlyPh	163
Db	468	AGTATGCTTTGGGGCCAAATGTGTATGTGTGTCTTGGCGAGGCTCATTCACGGGTT	527
QY	163	eGlyIleGlyLeuAlaValThrLeuValProLeuTyrIleSerGluIleAlaProSerG1	183
Db	528	CGGATCGGTTTGGCTGTACAGCTTGTACCATTTGTACATCTCTGAGACTGCCCGACGGA	587
QY	183	uIleArgGlyLeuLeuAsnThrLeuProGlnPheSerGlySerGlyGlyMetPheLeuSe	203
Db	588	CATCAGAGGACTGCTAAACAGCTGCCGAGTTCAGTGGGTCTGGAGGGATGTTCTTTTC	647
QY	203	rTyrCysMetValPheGlyMetSerLeuSerProSerProAspTrpArgIleMetLeuG1	223
Db	648	ATACTGCATGGTATTGGCATGTCCCTCATGCGCAGCCAGATTGGAGGATCATGCTTGG	707
QY	223	yValIleuAlaIleProSerLeuPhePhePheGlyLeuThrIlePheTyrLeuProGluSe	243
Db	708	CGTTCTATCAATACCATCACTTATATACCTTTGCTGATGACCATCTTTTACTTACCTGAATC	767
QY	243	rProArgTrpLeuValSerLysGlyArgMetAlaGluAlaLysLysValLeuGlnLysLe	263
Db	768	GCCGAGGTGGCTCGTGAGCAAAAGGAAGATGGCTGAGGCCAAGCGTGTGTGCAAGGCCT	827
QY	263	uArgGlyLysAspAspValSerGlyGluLeuSerLeuLeuGluGlyLeuGluValG1	283
Db	828	CGGTGAAGAGAGAGATGTTTCAGGAGAAATGSCCTTCTCGTTGAAGGTCTGGGGTGG	887
QY	283	yGlyAspThrSerIleGluGluTyrIleIleGlyProAlaThrGluAlaAlaAspAspLe	303
Db	888	GAAAGACACAAAATTGAGGAATACATAATTGACCTGATGATGAGCTGTGATGAGG	947
QY	303	uValThrAspGlyAspLysGluGlnIleThrLeuTyrGlyProGluGlyGlnSerTr	323
Db	948	GCTGGCTCCA--GATCCAGAGAGAGATCAAACTGTATGGTCTCTGGAAGAGGCTTATCGTG	1004
QY	323	pIleAlaArgProSerLysGlyProIleMetLeuGlySerValLeuSerLeuAlaSerAr	343
Db	1005	GGTTGCCCGTCTCTGTTCAAGGCAAGTGGCTGGAAGTGCATTAGGTCTCATCTCTCG	1064
QY	343	gHisGlySerMetValAsnGlnSerValProLeuMetAspProIleValThrLeuPheG1	363

Db 1065 TCATGGTAGTATGGTCAGTCAGGGTAAGCCCTTGTGGATCCTGTTGTCAACCCTTTTGG 1124

QY 363 ySerValHisGluAsnMetProGlnAlaGlyGlySerMetArgSerThrLeuPheProAs 383

Db 1125 AAGTCTCCATGAGAAGATGCCTGAGATAATGGGAAGCATGCGGAGCACATTGTTTCCTAA 1184

QY 383 nPheGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTrpAspGluG1 403

Db 1185 CTTTGGCAGCATGTTTAGTGTGSCGGAACAGCAGCAAGCTAAAGGTGATTGGGATGCTGA 1244

QY 403 uAsnLeuHisArgAspAspGluGluTyrAlaSerAspGlyAlaGlyGlyAspTyrGluAs 423

Db 1245 GAGT---CAACGGGAGGGTGAAGATTATGGATCAGACCATGGTGGGATGACATTGAAGA 1301

QY 423 pAsnLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIle-- 442

Db 1302 TAGCCTCCAAAGCCCACTTATTCTCGTCAAGCGACAAGCGTGAAGGAAGGAGATCGC 1361

QY 443 -valHisHisGlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeuGlyG1 462

Db 1362 TGCACCTCATGGCAGTATAATGGTGTCTGTGGGA-----AGAAAGTAGTAGTCTCATGCA 1415

QY 462 uGlyGlyAspGlyValSerSerThrAspIleGlyGlyTrpGlnLeuAlaTrpLysTr 482

Db 1416 GGGCGGAGGCAGTAAGCAGCATGGGCATTGGTGGGGATGGCAGTTGGCTTGGAAATG 1475

QY 482 pSerGluLysGluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyrLeuHi 502

Db 1476 GACTGAGAGAGAAGGTGCAGATGGCGAAAAAGAGGTGGCTTCCAACGTATCTACTTCCA 1535

QY 502 sGlnGluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGlyGlyAs 522

Db 1536 TGAAGAGGGTGTGACAGGTGATCGCAGGGGCTCTATACTGTGCATTGCCT---GGAGGTGA 1592

QY 522 pValPheGluGlySerGluPheValHisAlaAlaLeuValSerGlnSerAlaLeuPh 542

Db 1593 TGTTCTCTCTGGTGGTGAGTTCGTCCAGGCAGCTGCTCTGTCCAGCCAACCTGTCTCTTA 1652

QY 542 eSerLysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisProSerGluVa 562

Db 1653 CTCTAAGGAATTGATGAAGCAACGCCTGTGCTGCCCTGCTATGTCATCCATCTCAGGC 1712

QY 562 lAlaAlaLysGlySerArgTrpLysAspLeuPheGluProGlyValArgArgAlaLeuLe 582

Db 1713 AGTTGCTAAAGGTCCTAAATGGGCAGACTTATTGCAACCTGGAGTGAAGCATGCTCTGTT 1772

QY 582 uValGlyValGlyIleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyValLeuTyrTy 602

Db 1773 TGTTGGCATAGGGATACAAATCCTGCAACAGTTTGTCTGGCATTATGGAGTTCTGTACTA 1832

QY 602 rThrProGlnIleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSe 622

Db 1833 CACTCCACAATACTTTGAGCAAGCTGGTGTGGTGTCTTCTTGCAACATTGGACTTAG 1892

QY 622 rSerAlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIl 642

Db 1893 CTCCTCATCTGCATCTATTCTTATTAGCGGACTGACAAACCTTGCTGACGCTTCCCAGCAT 1952

QY 642 eGlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuGlyThrIl 662

Db 1953 TGGTATTGCTATGAGGCTCATGGATATGCTGGAAGAAGGTTTCTTCTCCTTGCAACAAT 2012

QY 662 eProIleLeuIleAlaSerLeuValIleLeuValValSerAsnLeuIleAspLeuGlyTh 682

Db 2013 CCCTATCCTGATAGTACACTAGCTATCTTGTATCTGTGATATCTGTCATATCTGGGAC 2072

QY 682 rLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyrPheCysCysPheValMe 702

Db 2073 CATGGTTTCATGCTCACTGTCCACAGTCAGTGTCTACTCTACTTCTGCTTCTTTGTGAT 2132

QY 702 tGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePheProThrArgValArgGl 722

Db 2133 GGGGTTCCGGGCCTATTCCAAACATTCTCTGTGCAGAGATTTTCCCGACCAACCGTTCCGTGG 2192

QY -722 yLeuCysIleAlaIleCysAlaPheThrPheTrpIleGlyAspIleIleValThrTyrSe 742

Db 2193 CATCTGCATAGCCATCTGTGCCCTAACATTCTGGATCGGTGATATCATTTGTGACATACAC 2252

QY 742 rLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValVa 762

Db 2253 CCTCCCCGTGATGCTCAACGCCATTGGACTCGCTGGAGTGTGGAATCTACGCAGTGGT 2312

QY 762 lCysLeuIleSerPheValPheValPheLeuLysValProGluThrLysGlyMetProLe 782

Db 2313 CTGCATACCTGGCTTCTCTGTTTGTCTTCATGAAGGTGCCGGAGACAAAGGCGATGCCCTCT 2372

QY 782 uGluValIleThrGluPhePheAlaValGlyAlaLysGlnAla 796

Db 2373 TGAAGTCATCACCGAGTCTTCTCTGTCGGAGCAAAGCAGGCC 2415

RESULT 7

HVU534445

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

HVU534445 2614 bp mRNA linear PLN 14-FEB-2003
Hordeum vulgare mRNA for hexose transporter (stp1 gene).

AJ534445 GI:26986185
hexose transporter; stp1 gene.
Hordeum vulgare subsp. vulgare
Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooideae; Triticeae; Hordeum.

Weschke,W., Panitz,R., Gubatz,S., Wang,Q., Radchuk,R., Weber,H. and
Wobus,U.
The role of invertases and hexose transporters in controlling sugar
ratios in maternal and filial tissues of barley caryopses during
early development

Plant J. 33 (2), 395-411 (2003)

22424051

12535352

2 (bases 1 to 2614)

Radchuk,R.

Direct Submission

Submitted (09-DEC-2002) Radchuk R., Molecular Genetics - Gene
Expression, Plant Genetics and Crop Plant Research, Corrensstr. 3,
Gatersleben, D-06466, GERMANY

Location/Qualifiers

1. .2614

/organism="Hordeum vulgare subsp. vulgare"

/mol_type="mRNA"

/cultivar="barke"

/sub_species="vulgare"

/db_xref="taxon:112509"

/tissue_type="developing caryopsis"

/clone_lib="lambda ZAP2"

/dev_stage="0-15 days after flowering"

1. .2614

/gene="stp1"

83. .2314

/gene="stp1"

/codon_start=1

/product="hexose transporter"

/protein_id="CAD58958.1"

/db_xref="GI:26986186"

/db_xref="GOA:Q8GT52"

/db_xref="TrEMBL:Q8GT52"

/translation="MSSGAVLVVAIAASIGNLLQGWDNATIAGAVLYIKKEFSLETQPL
IEGLIVAMSLIGATVITTFSGAVADAVGRRPLLIASSVLYFVSGLVMLWAPNVYVLL
ARLDGFIGLAVTLVPLYISETAPTDIRGLNLTLPQFSGSGMFLSYCMVFTMSLMP
OPDWRIMGLVLSIPSLMYFALTVPYLPESPRWLVSCKRMABAKRVLQRLRGREDVSGE
MALLVEGLVGKDTHEEYIIGPDDELADGLAPDQEKLYGAEVGSWIARPVXG
GQSALGSALGLMSRHGSMVSGKSLVPLVTLFSGVHEKXMEVMSGMRSTLFFNFGSM
FSVAEQQAKADWDAESHDDDEDYASDHGADDIEDNLSLISQATSVEGKEIAAPH
GSIIMGVESSMQGGDAVSSMGIGGWQLAWKTWEREGADQKEGGFORIYLHEEVS
GDRRGSIILSMPPGGDIPPGGEYIQAAALVSQALYSKDLEEQQLAGPAMVHPSEAVAK

TKWAELFEPGVKHALFVIGLQILOQFAGINGVLYYTPQILEQAGVGILLNSIGLSS
SASILISALTTLMLPSIGIAMRLMDMSRRFLLLSTIPVLIVALAILVINVLDVGT
MVHAALSTISIVYFCFFVMGFGPIPNILCAEIPFTSVRGICAIICALTFWIGDIIVT
YTLPLVNLNAIGLAGVGIYAVVCMIAFVYVMKVPETKGMPLVITEFFSVGAKQKE
ATD"

ORIGIN

Alignment Scores:
Pred. No.: 7.2e-178 Length: 2614
Score: 2808.00 Matches: 543
Percent Similarity: 84.39% Conservative: 84
Best Local Similarity: 73.08% Mismatches: 104
Query Match: 68.91% Indels: 12
DB: 8 Gaps: 6

US-10-051-909-32 (1-800) x HVU534445 (1-2614)

Qy	57	SerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrpAsp	76
Db	89	TCGGCGCCCGTGTGGTCGCATAGCGCGTCCATCGGAAACCTGTCTGACGGGTGGAT	148
Qy	77	AsnAlaThrIleAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsnGlu	96
Db	149	AATCGACCATTCAGGTGCGGTCTGTACATAAAGAGGAGTTTCAGCCTGGAGACCCAG	208
Qy	97	ProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr	116
Db	209	CCCTTGATCGAGGCGCTCATCGTGGCCATGTCGTCATCGGAGCGACGGTTATCACGACG	268
Qy	117	PheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSerIle	136
Db	269	TTCTCCGGGGCGGTGGCTGACGCTGTTGGTAGGGGCCCTGCTGATCGCTCGTCTGTC	328
Qy	137	LeuTyrPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyrValLeuLeu	156
Db	329	CTCTACTTTGTCACTGGCGCTGGTGATGCTCTGGCGGCCCAACGCTCTATGTTGCTCTTG	388
Qy	157	AlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyrIle	176
Db	389	GCCAGGCTCATCGACGGGTTCGGTATCGGTTTGCAGTCAACCTTGTCTCTGTACATT	448
Qy	177	SerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSerGly	196
Db	449	TCGGAGACCGCCCGACTGACATTAGAGGGCTGCTGAACACGCTGCCGAGTTTCAGTGGG	508
Qy	197	SerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSerPro	216
Db	509	TCAGGAGGATGTTCTTCTTACTGATGGTGTATTACCATGTCTCCCTCATGCCGAGCCT	568
Qy	217	AspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeuThr	236
Db	569	GACTGGAGAATCATGCTGGGGTTTGTTCGATCCCGTCGCTTATGATTGTCATTGACT	628
Qy	237	IlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGluAla	256
Db	629	GTCITTTATTGGCCGAGTCGCCAGATGGCTTGTGAGCAAGGAGAAATGGCTGAGGCC	688
Qy	257	LysLysValLeuGlnLysLeuArgGlyLysAspValSerGlyGluLeuSerLeuLeu	276
Db	689	AAGCGAGTATTGCAAGACTGCGGGGAGGGAAGATGCTCTCAGGAGATGGCCCTTCTT	748
Qy	277	LeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIleGlyProAla	296
Db	749	GTTGAAGGATTGGGTGTGGAAAGACACATTTTGAGGAATACATAATTGGACCTGAT	808
Qy	297	ThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyrGly	316
Db	809	GATGAGCTGGTGTGATGACGGGCTGGCTCCA---GATCAAGAGAAGTTGAAACTCTATGGA	865
Qy	317	ProGluGluGlyGlnSerTrpIleAlaArgProSerLys-----GlyProIleMetLeu	334
Db	866	GCTGAAGAAGGGGTATCTTGGATCGCCCGTCTGTTAGGSGTGGCGGCCCAAGTGCACTT	925

Qy	335	GlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValProLeu	354
Db	926	GGAAAGCGCTTTGGCGCTCATGTCTCGTCATGGAGTATGTTAGTCAGGGTAAATCTCTT	985
Qy	355	MetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGlyGly	374
Db	986	GTTGACCCCTTGTCACTCTCTTCGGAAGTGTCCATGAGAAAGATGCCGTAGTAATGGGG	1045
Qy	375	SerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGlnHis	394
Db	1046	AGCATGCGGAGCACCTTGTTCCTAACTTTGGCAGCATGTTTAGTGTGGCTGAACAGCAG	1105
Qy	395	AlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspGluGluTyrAlaSer	414
Db	1106	CAGGCTAAAGCTGACTGGGATGCTGAGAGT---CATAGGATGATGAAGATTATGCATCG	1162
Qy	415	AspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAla	434
Db	1163	GATCATGGTGTCTGATGACATTGAGGATAACCTCAATAGCCCGCTTATTTCTCGTCAAGCG	1222
Qy	435	ThrGlyAlaGluGlyLysAspIleVal-----HisHisGlyHisArgGlySerAlaLeu	452
Db	1223	ACAAGCGTGGAGGGTAAGAGATTGCTGCACCTCATGAAGCATATAATGGGTGGTGGAA	1282
Qy	453	SerMetArgArgGlnThrLeuLeuGlyGluGlyGlyValSerSerThrAspIle	472
Db	1283	AGTAGTAGCATGCAG-----GGAGGGGACGCAGTAAGCAGCATGGGCATT	1327
Qy	473	GlyGlyGlyTrpGlnLeuAlaTrpLysTrpSerGlnLysGluGlyGluAsnGlyArgLys	492
Db	1328	GGTGGGGGTGGCAGTTAGCTTGGAAAGTGGACTGAGAGAGAGGTGCACATGGGCAAAA	1387
Qy	493	GluGlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArgGly	512
Db	1388	GAAGTGGCTTCCAGCGTATTACTTGCATGAGGAGGGTGTGTCAAGTGCATCGGAGAGGC	1447
Qy	513	SerIleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAla	532
Db	1448	TCTATATTGTTATGCCA---GGAGGTGATATTCTCTCTGGTGGTGGTATATCCAGGCA	1504
Qy	533	AlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSer	552
Db	1505	GCCGCTCTAGTGAGCCAACTGCTCTTTACTCGAAGGACCTGTATAGACGACGCTTGCT	1564
Qy	553	AspAlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeu	572
Db	1565	GGTCCAGCCATGTTACATCCATCCGAGGAGTGTCCAAAGGGTACAAAGTGGGCAGAACTA	1624
Qy	573	PheGluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnGln	592
Db	1625	TTTGAACCTGGAGTGAAGCATGCATGTTTGTGGCATTTGGATTACAGATCTCTGCAGCAG	1684
Qy	593	PheAlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluAlaGlyVal	612
Db	1685	TTTGCGGTATCAACGGAGTCTCTGTATTACACGCTCAGATACCTTGAGCAAGCAGGTGT	1744
Qy	613	AlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSer	632
Db	1745	GGGATTTCTTCTATCAAAACATTGGACTAAGCTCTCTCTGTCATCTATTCTTATTAGTGCC	1804
Qy	633	LeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSer	652
Db	1805	TTGACAACTTGTGTGATGCTTCCAGCATTTGGCATCGCCATGAGACTCATGGATATGTCA	1864
Qy	653	GlyArgArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeu	672
Db	1865	GGAAGAAGGTTTCTCTCTCTCAACATCCCTGTCTTGATAGTGGCACTAGCTATCTTG	1924
Qy	673	ValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSer	692
Db	1925	GTTTTGGTGAATGTTCTGGATGTGGGAACCATGGTGCACGCCGCTCTCTCAACGATCAGT	1984
Qy	693	ValIleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCys	712

Db 1985 GTCATAGTCTACTTCTGCTTCTTCGTCATGGGTTTGGACCTATCCCGAATATTTCTGTC 2044

Qy 713 AlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPhe 732

Db 2045 GCGGAGATTTTCCCACTCCGTCGTCGTCATCTGCAATGCCATCTGCGCACTAACCTTC 2104

Qy 733 TrpIleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeu 752

Db 2105 TGGATTGGTGACATCATCGTGACATACACTCTTCCCGTGATGCTCAACGCCATCGGACTC 2164

Qy 753 AlaGlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeu 772

Db 2165 GCTGGAGTTTTCGGAATATATGCTGCTGCTTGGCATGATACCTTCGTATTGCTACATG 2224

Qy 773 LysValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGly 792

Db 2225 AAGGTCCTGAGACAAAGGGCATGCCCTGGAGGTCATCACCGAGTTCTTCTCGGTGGG 2284

Qy 793 AlaLysGln 795

Db 2285 GCGAAGCAG 2293

RESULT 8

AR208568

LOCUS AR208568 2601 bp DNA linear PAT 20-JUN-2002

DEFINITION Sequence 7 from patent US 5383776.

ACCESSION AR208568

VERSION AR208568.1 GI:21509755

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 2601)

AUTHORS Allen,S.M., Hitz,W.D., Kinney,A.J. and Tingey,S.V.

TITLE Plant sugar transport proteins

JOURNAL Patent: US 5383776-A 7 07-MAY-2002;

FEATURES

source

1..2601

/organism="unknown"

/mol_type="unassigned DNA"

ORIGIN

Alignment Scores:

Pred. No.: 6.32e-169 Length: 2601

Score: 2674.00 Matches: 522

Percent Similarity: 82.20% Conservative: 92

Best local Similarity: 69.88% Mismatches: 115

Query Match: 65.62% Indels: 18

DB: 6 Gaps: 8

US-10-051-909-32 (1-800) x AR208568 (1-2601)

Qy 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrp 75

Db 175 ATGAAAGGTGCGGTCCTTGTGCTATTGCGGCTTCCATTGGTAAATTCCTCCAAGGATGG 234

Qy 76 AspAsnAlaThrIleAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsn 95

Db 235 GATAATGCTACCATCGCCGGGCTAATGGTTACATTAAAGAAAGACCTTGCTTTGGGAACA 294

Qy 96 GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115

Db 295 -----ACTATGGAAGGCTTGTGGTGGCATGTCCCTGATTGGAGCAACGGTAATCACCC 348

Qy 116 ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer 135

Db 349 ACATGCTCTGCTCTATAGCGGATTGGTCGGTCGGCGACCCATGATGATAATCTCATCT 408

Qy 136 IleLeuTyrPhePheSerGlyLeuIleMetLeuTyrSerProAsnValTyrValLeuLeu 155

Db 409 GTGCTCTATTCTTGGGTGGTTTGGTGATGCTGTGGTCCCCCAATGTGTATGTGTGTC 468

Qy 156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr 175

Db 469 TTGGCGAGGCTACTTGTATGGATTTGGATTGGCTTGTCTGTGACTTGTGCCGGTCTAT 528

Qy 176 IleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSer 195

Db 529 ATATCTGAACGGCGCGCTCTGAATAAGGGGGTCGTTGAATACGCTTCTCAGTTCACT 588

Qy 196 GlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSer 215

Db 589 GGCTCTGGAGGAATGTTTTCGFACTGTATGGTTTGGCATGTCTATTTAGTCCCGCG 648

Qy 216 ProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeu 235

Db 649 CCTAGCTGGAGGCTCATGCTTGGGTTCTGTCTATTCTCTCTCTTGTATTGTGCATTG 708

Qy 236 ThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGlu 255

Db 709 ACCATTTTCTTGTCCCGAGTCTCTCTCGGTGGCTGGTCAGCAAGAGGATGCTCGAG 768

Qy 256 AlaLysLysValLeuGlnLysLeuArgGlyLysAspValSerGlyGluLeuSerLeu 275

Db 769 GCTAAGAAGGTGCTCCAAAGATTGCGGGAAGGAGGATGTGTGAGCGGAGATGGCATTG 828

Qy 276 LeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIleGlyPro 295

Db 829 CTGGTTGAAGGCTCTCGGATTGGGGTGATACATCTATCGAAGAGTACATAATTGGCCT 888

Qy 296 AlaThrGluAlaAlaAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyr 315

Db 889 GCTGACGATGTGGCTGATGGTCATGATGATGCAACAGAGAGATAAAATTCGATTATAT 948

Qy 316 GlyProGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGly 335

Db 949 GGATCCCAAGCAGGCTTCTTCTGGTTATCAAAACCTGTCACTGGACAGAGTTCTATTGGC 1008

Qy 336 SerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValProLeuMet 355

Db 1009 -----CTTGGCTCACACCATGGAAGCATCATCAACCAAGCATGCCCTCATG 1056

Qy 356 AspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly----- 373

Db 1057 GATCCTCTGGTGACACTGTTTGGTAGCATTCATGAGAAGCTCCCCCGAGACAGGACAAGA 1116

Qy 374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393

Db 1117 GGAAGCATGCGAAGCACTCTGTTTCCAAATTTTGGAAGCATGTTCAAGCACTGCTGAGCCG 1176

Qy 394 HisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAla 413

Db 1177 CATGCTAAATTCGAACAATGGGATGAGAAAGCTTACAAAGGGAACCTGAGGACTACATG 1236

Qy 414 SerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGln 433

Db 1237 TCAGATGCAACCCGTGGGACTCCGATGATAAATTTGCACAGTCCCTTAATCTCACGCCAA 1296

Qy 434 AlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeu--- 452

Db 1297 ACAACAAGCCTTGAA---AAAGACTTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTTGGC 1353

Qy 453 SerMetArgArgGlnThrLeuLeuGlyGluGly---GlyAspGlyValSerSerThrAsp 471

Db 1354 AGCATGAGCGGTCACAGTAGTCTCATGCAAGGTCAGGTGAGCAAGGTGAGTACAGGT 1413

Qy 472 IleGlyGlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArg 491

Db 1414 ATTGGTGGTGGCTGGCAACTGGCATGGAAATGGACTGATAAA---GGTGAGGATGGAAA 1470

Qy 492 LysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArg 511

Db 1471 CAACAAGGAGGGTTAAAGGATTTATTTATACATGAGGAGGGAGTTTCTGCATCTCGTCGT 1530

Qy 512 GlySerIleValSerLeuProGlyGlyAspValPheGluGlySerGluPheValHis 531

Db	1531	GGATCCATTGTATCGATTCCCGGTGAAGGC-----GAATTTGTCCAG	1572
Qy	532	AlaAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMet	551
Db	1573	GCTGCTGCCTTGGTAAGCAACCCCGCTCTTTACTCAAGGAGCTTATTGATGGACACCCA	1632
Qy	552	SerAspAlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAsp	571
Db	1633	GTTGGCCCTGCAATGGTTCAACCCATCTGAGACAGCTTCAAGGGGCCAAGTTGGAAGCT	1692
Qy	572	LeuPheGluProGlyValArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGln	591
Db	1693	CTTCTTGAACCAAGGGGTTAAGCATGCATTGGTTGTTGGAGTTGGAATACAAATACTTCAG	1752
Qy	592	GlnPheAlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGly	611
Db	1753	CAGTTTTCAGGGATAAATGGGTTCTATATTACACACCTCAAAATCCTTGAAGAGCCCGGT	1812
Qy	612	ValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSer	631
Db	1813	GTTGAAGTTCTTCTTTCAGATATAGGCATTGGCTCAGAGTCGGCATCATCTTATCAGT	1872
Qy	632	SerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeu	651
Db	1873	GCTTTCACAACCTTCTTGATGCTTCCCTGTATAGCGGTAGCCATGAAGTCATGGATGT	1932
Qy	652	SerGlyArgArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIle	671
Db	1933	TCAGGCAGAAGGCAGTTGCTACTTACTACAATCCCCGCTGCTGATTGTGTCATCTATT	1992
Qy	672	LeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrVal	691
Db	1993	TGCTCATTCGAAGCCCTGGTAAATTTTGGCAATGTGCCCCATGCAGCAATCTCAACAGTA	2052
Qy	692	SerValIleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeu	711
Db	2053	TGCGTTGTGTTTATTCTGCTGCTTTGTGATGGTTATGGACCAATTCCAAACATCCTT	2112
Qy	712	CysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThr	731
Db	2113	TGCTCAGAGATTTTCCCACTAGGTGCGTGCGCTCTGCATTGCTATCTGTGCATTAGTG	2172
Qy	732	PheTrpIleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGly	751
Db	2173	TTCTGGATTGGAGACATCATCATCATCTCGCTGCCTGTGATGCTCGGCTCTTTAGGA	2232
Qy	752	LeuAlaGlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPhe	771
Db	2233	CTTGGTGGTGTATTCCGCATTTACGCAGTTGTGTTTCTCATCTCGTGATATTGTGTTT	2292
Qy	772	LeuLysValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaVal	791
Db	2293	TTGAAGGTTCCAGAAACAAAGGCGCATGCCCTTGAAGTCATCTCTGAATCTTTTCTGTT	2352
Qy	792	GlyAlaLysGlnAlaAlaAla 798	
Db	2353	GGAGCAAGCAGGCTGCTTCT 2373	
RESULT	9		
AY094465			
LOCUS	AY094465	2570 bp mRNA linear PLN 05-MAY-2002	
DEFINITION	Arabidopsis thaliana AT4935300/F23E12_140 mRNA, complete cds.		
ACCESSION	AY094465		
VERSION	AY094465.1	GI:20453188	
KEYWORDS	FLI CDNA.		
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;		
	rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.		
REFERENCE	1 (bases 1 to 2570)		
AUTHORS	Shinn,P., Chen,H., Cheuk,R., Kim,C.J., Meyers,M.C., Banh,J.,		

TITLE	Arabidopsis cDNA clones		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 2570)		
AUTHORS	Shinn,P., Chen,H., Cheuk,R., Kim,C.J., Meyers,M.C., Banh,J., Bowser,L., Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.		
TITLE	Arabidopsis cDNA clones		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 2570)		
AUTHORS	Shinn,P., Chen,H., Cheuk,R., Kim,C.J., Meyers,M.C., Banh,J., Bowser,L., Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.		
COMMENT	RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-length cDNA') : Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.		
FEATURES	Location/Qualifiers		
source	1. 2570		
	/organism="Arabidopsis thaliana"		
	/mol_type="mRNA"		
	/db_xref="taxon:3702"		
	/chromosome="4"		
	/clone="RAFL09-48-J19 (R21968)"		
	/ecotype="Columbia"		
	209..2428		
CDS	/note="putative sugar transporter protein"		
	/codon_start=1		
	/product="AT4g35300/F23E12_140"		
	/protein_id="AAM19835.1"		
	/db_xref="GI:20453189"		
	/translation="MSGAVLVAIAAAVGNLLQWDNATAGAVLYIKKEFNLESNPV EGLIVAMSLIGATLITTCGGVADWLGRRPMLILSSILYFVGLVLMWSPNVYVLLLG RLDDGFGVGLVTLVPIYISETAPPEIRGLNLTLPQFTSGGMFLSYCMVFGMSLMP PSWRLMLGLVFIPLVFFFLTVFFLPESPRWLVSCKRMLEAKRVLQRLRGREDVSGEM ALLVEGLGIGGETTIEEYIIIPADEVTDDHDIAVDKQIKLYGAEGLSWARPVKGG STMSVLSRHGSTMRRQSGSLIDPLVTLFGSVHEKMPDTGSMRSALFPFHGSMFVGGN QPRHEDWDEENLVGEDYPSDHGDDSEDDLHSPILSRQTSMEKOMPHTAHGTLSTF RHGSVQQAQGEAGSMGIGGWQVANKWTEREDESQKEGGFKRIYILHQEGFPGSRR GSIIVSLPGDGTGEADFVQASALVQPALYSKDLKEHTIGPAMVHPSETTKGSIWHD LHPGVKRALVVGVLQILQQFSGINGVLYYTPQILEQAGVILLNMGISSSASLL ISALTTFVMLPAIAVAMRLMDLSGRRTLLTTPILIASLLVLSNLMVHNSIVHAV LSTVSVLYFCFFVMGFGPAPNLICSEIFPTRVRGICIAICALTFCIDIIIVTSLPV LLKISIGLAGVFGMYAIVCCISNVFVFIKVPETKMPLEVIETFFSVGARQAEAAKNE"		
ORIGIN			
Alignment Scores:	8.48e-168	Length:	2570
Pred. No.:	2657.00	Matches:	530
Score:			

Percent Similarity: 79.87%		Conservative: 97	
Best Local Similarity: 67.52%		Mismatches: 133	
Query Match: 65.20%		Indels: 25	
DB: 8		Gaps: 10	
US-10-051-909-32 (1-800) x AY094465 (1-2570)			
QY	29	LeuAlaLeuProGlyProLeuProAlaSerCysSerSerGlnGlu-----	44
Db	91	CTTCTTCTGCCT--CCAATCCAAGTGGATTTGTTTCACTCACTCATCGTGTGATC	147
QY	45	-----ProValThrSerAspAspIleLeuGluAspLys-----	55
Db	148	AGTGACTCTGTTTTTTCGAAATTCCTCTCAGATTTCTTGATAAAAGATAGAACGGAAT	207
QY	56	-MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTr	75
Db	208	TATGAGTGGAGCTGTGCTTGTGCTATTGCTGCTGCTGTTGGCAACTTGTACAGGATG	267
QY	75	pAspAsnAlaThrIleAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAs	95
Db	268	GGATAACGCAACTATTGCAGGAGCTGTGTGTACATAAAAAAGGAGTTTAAATTGGAGAG	327
QY	95	nGluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValTh	115
Db	328	TAATCCATCAGTGAAGGTCTAATTGTGGCGATGTCACTTATTGGTGCTACTCTGATTAC	387
QY	115	rThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSe	135
Db	388	AACATGCTCTGGAGGGGTAGCTGATTGGCTTGGTCGCCGTCCCATGCTAATAATTGTCCTC	447
QY	135	rIleLeuTyrPhePheSerGlyLeuIleMetLeuTyrSerProAsnValTyrValLeuLe	155
Db	448	AAITCTCTACTTTGTTGGTCTCTAGTAATGCTATGGTCTCCGAATGTTTATGTGTGCT	507
QY	155	uLeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTy	175
Db	508	CTTAGGAAGGTGTTAGATGGATTGGGGTTGGTCTGTGGTCACACTTGTTCCTATTTA	567
QY	175	rIleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSe	195
Db	568	TATATCTGAGACTGCACCCACTGAGATTAGGGGACTGTTGAATACGCTACCGCAGTTCAC	627
QY	195	rGlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSe	215
Db	628	TGGCTCTGGAGGGATGTTCTTATCTTACTGTATGGTTTTTCGGAATGTCGTTGATGCCATC	687
QY	215	rProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLe	235
Db	688	ACCTAGCTGGAGATTGATGCTTGGTGTCTCTTTTCATCCCTTCCCTTGTCTTTTCTTCCT	747
QY	235	uThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMetAlaG1	255
Db	748	CACGGTCTTCTTCTGCCCGAGTCCCAAGGTGGCTCGTGAGCAAAGGTGGAATGCTTGA	807
QY	255	uAlaLysLysValLeuGlnLysLeuArgGlyLysAspAspValSerGlyGluLeuSerLe	275
Db	808	AGCAAAGCGGGTCTTTCAGAGACTGCGTGGTCCGGAAGATGTGTCTGGTGAGATGGCTTT	867
QY	275	uLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIleGlyPr	295
Db	868	GTTGGTTGAGGGTCTTGGAAATTGGAGGTGAACAACCATAGAGGAATATATAATTGGTCC	927
QY	295	oAlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTy	315
Db	928	CGCGGATGAAGTTACTGATGATCATGATATAGCTGTGGATAAGGATCAAAATTAAGTTATA	987
QY	315	rGlyProGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuG1	335
Db	988	TGGTGCAGAGAAGGGCTGAGTTGGTTGCTAGGCCAGTCAAAAGGA-----GG	1035
QY	335	ySerValLeuSerLeuAlaSerArgHisGlySer---MetValAsnGlnSerValProLe	354

Db	1036	AAGCACTATGAGTGTTTTTGTCTGCCCATGGAAGTACAATCAGCAGGAGGCAAGGCTCAT	1095				
QY	354	uMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGlyG1	374				
Db	1096	GATTGATCCTCTTGTACACTGTTTGGGAGCGTTTCACGAGAAGATGCCGGACACT---GG	1152				
QY	374	ySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGlnHi	394				
Db	1153	AAGCATGAGGAGTGCTTGTTCACACATTTTGGGAGTATGTTCACTGTTGGAGGAATCA	1212				
QY	394	sAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAlaSe	414				
Db	1213	ACCAAGACATGAAGATTGGGATGAAGAGAATCTTGTGGAGAAGGTGAGGATTATCCATC	1272				
QY	414	rAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAl	434				
Db	1273	CGAC---CATGGAGATGATTCTGAAGATGATCTTCACTTCCGTTGATCTCACGTCAAAC	1329				
QY	434	aThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeuSerMe	454				
Db	1330	GACAAGCATGGAG---AAAGACATGCCTCACACTGCTCATGGAACCTTTTCTACCTTCAG	1386				
QY	454	tArgArgGlnThrLeuLeuGlyGluGlyGlyAspGlyValSerSerThrAspIleGlyG1	474				
Db	1387	ACATGGAAGTCAAGTGCAGGAGCTCAAGGGGAAGGAGCGGTAGTATGGGATTGGAGG	1446				
QY	474	yGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArgLysGluG1	494				
Db	1447	TGGATGGCAAGTGGCATGGAATGACGGAAGAGAGATGAATCGGCACAGAAAGAGG	1506				
QY	494	yGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArgGlySerI1	514				
Db	1507	TGGGTTTAAACGGATATACTTGCATCAAGAAGGTTTCCCAGGATCTCGACGTGGCTCAAT	1566				
QY	514	eValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaAlaA1	534				
Db	1567	TGTTTTCATTCCCTGGTGGTGTGTAACCGGTGAG--GCAGATTTGTACAAGCGTCTGC	1623				
QY	534	aLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAspAl	554				
Db	1624	TTTGGTTAGCCAAACAGCTCTTTATTCCAAAGACCTTCTCAAAGAACATACAATTGGTCC	1683				
QY	554	aAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuPheG1	574				
Db	1684	TGCTATGGTACATCCATCCGAA--ACAACTAAAGGGTCAATTGGCATGATCTTCATGA	1740				
QY	574	uProGlyValArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnGlnPheAl	594				
Db	1741	TCCTGGAGTCAAGCGTGCATTAGTCGTAGGAGTTGGACTTCAAATACATTCAGCAGTTC	1800				
QY	594	aglyIleAsnGlyValLeuTyrTrpThrProGlnIleLeuGluGlnAlaGlyValAlaVa	614				
Db	1801	AGGCATCAACGGAGTCTTTTACTACACACCGCAAAATCCTTGAGCAGCGGGTGTCTCGGAT	1860				
QY	614	lIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuTh	634				
Db	1861	CCTACTATCGAACATGGGATTAGTTCTTCTCCTCAGCATCCTTACTTATAAGTGCATTGAC	1920				
QY	634	rThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyAr	654				
Db	1921	AACCTTTGTGATGTACCTGCAATAGCTGTGCAATGAGGCTCATGGATCTTCTGGTGC	1980				
QY	654	gArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuValVa	674				
Db	1981	AAGGACCTTGCTTCTCACCCAGATACCAATCCTGTAGCATCTCTATTGGTTTAGTAAT	2040				
QY	674	lSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValI1	694				
Db	2041	CTCAAATCTTGTTCACATGAACAGCATTTGTGCACGGGCTCTTATCAACCGTAAGCGTGT	2100				
QY	694	eValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAlaG1	714				
Db	2101	GCTCTACTTCTGCTTCTTCGTGATGGGTTTCGGTCTCTGCTCCAAACATCTCTGTTTCA	2160				

156 LeuAlaArgPheValAspGlyPheGlyLeuAlaValThrLeuValProLeuTyr 175
 301 TTAGGAAGGTGTGTAGATGGATTGGGGTGGTCTTGGTCCACTTGTTCCTATTAT 360
 176 IleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSer 195
 361 ATATCTGAGACTGCACCACCTGAGATTAGGGGACTGTTGAATACGCTACCGCAGTTCACT 420
 196 GlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSer 215
 421 GGCTCTGGAGGGATGTTCTTATCTTACTGTATGGTTCGGAATGTCGTTGATGCCATCA 480
 216 ProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeu 235
 481 CCTAGCTGGAGATTGATGCTTGGTGTCTTTTCACTCCCTTCCCTTGTCTTTTCTTCTC 540
 236 ThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGlu 255
 541 ACCGCTCTTCTTCTCCCGAGTCCCAAGGTGGTCTCGTGAAGGTCGAATGCTTGAA 600
 256 AlaLysLysValLeuGlnLysLeuArgGlyLysAspValSerGlyGluLeuSerLeu 275
 601 GCAAAGCGGTCTTTCAGAGACTGCGTGGTGGGAGAGATGTCTGTTGAGATGGCTTGG 660
 276 LeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleGlyPro 295
 661 TTGCTTGAGGGTCTTGGAAATTGGAGGTGAACAACCATAGAGGAATATATATGTTCCC 720
 296 AlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyr 315
 721 GCGGATGAAGTTACTGATCATGATATAGCTGTGGATAGGATCAATTAAGTTATAT 780
 316 GlyProGluGluGlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGly 335
 781 GGTCCAGAAAGAGGGCTGAGTTGGTGTGCTAGGCCAGTCAAAGGA-----GGA 828
 336 SerValLeuSerLeuAlaSerArgHisGlySer---MetValAsnGlnSerValProLeu 354
 829 AGCACTATGAGTGTCTTCTCGCCATGGAAGTACAATGAGCAGGAGGCAAGGCTCATTTG 888
 355 MetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGlyGly 374
 889 ATTGATCCTCTTGTACACTGTTTGGAGCGTTCACGAGAAGATGCCGACACT---GGA 945
 375 SerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGlnHis 394
 946 AGCATGAGGAGTGCCTGTTTCCACATTTTGGAGTATGTTTCAAGTGTGGAGGGAATCAA 1005
 395 AlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAlaSer 414
 1006 CCAAGACATGAAGATTGGGATGAAGAGATCTGTTGGAGAAGTGCAGGATATCCATCC 1065
 415 AspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAla 434
 1066 GAC---CATGGAGATGATCTGAAGATGATCTTCAATCTCCGTTGATGCGGATTCACGTC 1122
 435 ThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeuSerMet 454
 1123 ACAAGCATGGAG---AAAGACATGCCTCACACTGCTCATGGAACCTTCTTCTACCTTCAGA 1179
 455 ArgArgGlnThrLeuLeuGlyGluGlyGlyAspGlyValSerSerThrAspIleGlyGly 474
 1180 CATGGAAGTCAAGTCGAGGAGCTCAAGGGGAAGGAGCGGGTAGTATGGGATTCGAGGT 1239
 475 GlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArgLysGluGly 494
 1240 GGATGGCAAGTCGATGGAATGGACCGGAAGAGAGATGAATCGGACAGAAAGAA--- 1296
 495 GlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgGlySerIle 514
 1297 -----GAGGTTTCCAGGATCTCGACGTGGCTCAATT 1329

515 ValSerLeuProGlyGlyGlyAspValPheGlyGlySerGluPheValHisAlaAlaAala 534
 1330 GTTTCATTCCTGGTGGTGGTGAACCGGTGAG---GCAGATTTTGTACAAAGCGTCTGCT 1386
 535 LeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAspAla 554
 1387 TTGGTTAGCCAAACAGCTCTTATTTCCAAAGACCTTCTCAAAGAACATACAATTGGTCTCT 1446
 555 AlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuPheGlu 574
 1447 GCTATGTTACATCCATCCGAA---ACAACTAAAGGGTCAATTTGGCATGATCTTCATGAT 1503
 575 ProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnPheAla 594
 1504 CCTGGAGTCAAGCGTGCATAGTCTGAGGAGTTGAGCTTCAAATACTTCAGCAGTTCTCA 1563
 595 GlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyValAlaVal 614
 1564 GGCATCAACCGGAGTCTTACTACACACCGCAATCCTTGAGCAGCGGGTGTCTCGGGATC 1623
 615 IleLeuSerLysPheGlyLeuSerSerAlaSerIleLeuIleSerSerLeuThr 634
 1624 CTACTATCGAAACATGGGAGTATGTTCTTCTCAGCATCCTTACTTATAAGTGCATTGACA 1683
 635 ThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArg 654
 1684 ACCTTGTGATGTTACCTGCAATAGTGTGTGCAATGAGGCTCATGGATCTTCTTGTGTCGA 1743
 655 ArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuValVal 674
 1744 AGGACCTTGTCTTCTCACCACGATACCAATCCTGTATGATCATCTCTATTGGTTTAGTAATC 1803
 675 SerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIle 694
 1804 TCAAATCTTGTTCACATGAACAGCATTTGTGCGCGGTCTTATCAACCGTAAGCGTTGTG 1863
 695 ValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAlaGlu 714
 1864 CTCTACTTCTGCTTCTCGTGTATGGGTTTCGTCCTGCTCCAAACATCCTCTGTTCCAGAG 1923
 715 IlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIle 734
 1924 ATTTTCCAACTCGAGTCGCGGAATCTGATCGCATCGCATCTGCGCACTCACCTTCTGGATC 1983
 735 GlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGly 754
 1984 TGTGACATAATCGTCACTTACAGTCTCCCGCTGCTGCTCAAATCCATTCGACTAGCTGGT 2043
 755 ValPheSerIleThrAlaValValCysLeuIleSerPheValPheValPheLeuLysVal 774
 2044 GTGTTTGAATGTACGCAATCGTATGTTGCAATTCATGGGTCTTTTGTGTTCAATAAGTC 2103
 775 ProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAlaLys 794
 2104 CCGGAACTAAAGCATGCCACTTGAAGTCAATCATCAGAGTCTTTTCTGTTGGAGCTAGA 2163
 795 GlnAlaAlaAla 798
 2164 CAAGCTGAAGCT 2175

RESULT 12

AK065191

LOCUS

DEFINITION

insert sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AK065191 2368 bp mRNA linear PLN 24-JUL-2003
 Oryza sativa (japonica cultivar-group) cDNA clone:J013002E10, full

Oryza sativa (japonica cultivar-group)
 AK065191.1 GI:32975209
 FLI_CDNA; CAP trapper.
 Oryza sativa (japonica cultivar-group)
 Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzae; Oryza.

REFERENCE
AUTHORS

1 The Rice Full-length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group: Otomo, Y., Murakami, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN: Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M., and Hayashizaki, Y.

TITLE

Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice

JOURNAL

Science 301 (5631), 376-379 (2003)

MEDLINE

22752273

PUBMED

12863764

REFERENCE

2 (bases 1 to 2368)

AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Namiki, T., Narikawa, R., Niikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Osato, N., Ota, Y., Otomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.

TITLE

Direct Submission

JOURNAL

Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp, Tel: 81-29-838-7007, Fax: 81-29-838-7007)

COMMENT

This clone is one of the 28K full-length cDNA clones from japonica rice.

URL : <http://cdna01.dna.affrc.go.jp/cdna/>

NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and Yamamoto, M.

FAIS Genome Sequencing & Analysis Group: Otomo, Y., Iida, Y.,

Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M.,

Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J.,

Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S.,

Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S.,

Yoshimura, A., Matsubara, K. and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K.,

Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T.,

Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K.,

Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y.,

Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J.,

Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,

Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,

Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.

FEATURES

source

1. 2368

/organism="Oryza sativa (japonica cultivar-group)"

/mol_type="mRNA"

/cultivar="Nipponbare"

/db_xref="taxon:39947"

/clone="J013002E10"

ORIGIN

Alignment Scores:
Pred. No.: 6.58e-163 Length: 2368
Score: 2583.00 Matches: 508
Percent Similarity: 83.74% Conservative: 69
Best Local Similarity: 73.73% Mismatches: 106
Query Match: 63.39% Indels: 7
DB: 8 Gaps: 5

US-10-051-909-32 (1-800) x AK065191 (1-2368)

QY 109 IleGlyAlaThrIleValThrPheSerGlyProLeuSerAspSerIleGlyArgArg 128
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 2 CTTGGGGCGACGATCATCACGAGTCTCTGGAGCAGTGGCTGATTTCTTTTGGTAGCGG 61
QY 129 ProMetLeuIleLeuSerSerIleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSer 148
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 62 CCCATGCTGATCGCGTGGCTGCTCTACTTTTGTAGTGGCTAGTGAATGCTTGGCG 121
QY 149 ProAsnValTyrValLeuLeuAlaAargPheValAspGlyPheGlyIleGlyLeuAla 168
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 122 CCAATGTGTATGTGTGCTCTTGGCGAGGCTCATTCAGGGGTTCGGGATCGGTTGGCT 181
QY 169 ValThrLeuValProLeuTyrIleSerGluIleAlaProSerGluIleAargGlyLeu 188
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 182 GTCACGCTGTACCATTTGTATCTCTGAGACTGCCCCGACGACATCAGAGGACTGCTA 241
QY 189 AsnThrLeuProGlnPheSerGlySerGlyGlyMetPheLeuSerTyrCysMetValPhe 208
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 242 AACACGCTCCGCGAGTTCAGTGGGTCTGGAGGGATGTCTTTCATCTACTGATGTTT 301
QY 209 GlyMetSerLeuSerProSerProAspTrpArgIleMetLeuGlyValLeuAlaIlePro 228
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 302 GGCATGTCCTCATGCCACAGCCAGATTTGGAGGATCATGCTTGGCGTTCTATCAATACCA 361
QY 229 SerLeuPhePhePheGlyLeuThrIlePheTyrLeuProGluSerProArgTrpLeuVal 248
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 362 TCACCTATATACCTTTGCATTGACCATCTTTTACTTACCTGAATCGCCGAGTGGCTCGTG 421
QY 249 SerLysGlyArgMetAlaGluAlaLysLysValLeuGlnLysLeuArgGlyLysAsp 268
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 422 AGCAAGGAAGAATGGCTGAGGCCAAGCGTGTGTGCAAGGCCCTGCGTGGAGAGAGAT 481
QY 269 ValSerGlyGluLeuSerLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIle 288
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 482 GTTTCAGGAGAAATGGCCCTCTCGTTGAAGGTCTGGGGGTGGGAAAGACACAAAATT 541
QY 289 GluGluTyrIleIleGlyProAlaThrGluAlaAlaAspAspLeuValThrAspGlyAsp 308
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 542 GAGGAATACATAATTGGACCTGATGATGAGCTTGTGCTGATGAAGGGCTGGCTCCA---GAT 598
QY 309 LysGluGlnIleThrLeuTyrGlyProGluGluGlyGlnSerTrpIleAlaAargProSer 328
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 599 CCAGAGAAGATCAAACTGTATGTTCTCTGAAGAGAGGCTTATCGTGGTGGCGCTCTGTT 658
QY 329 LysGlyProIleMetLeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetVal 348
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 659 CACGGGCAAAAGTGCACTTGGGAAGTGCATTAGGTCTCTCTCTCTCTCTCTCTCTCTCTCT 718

QY 349 AsnGlnSerValProLeuMetAspProIleValThrLeuPheGlySerValHisGluAsn 368
 Db 719 AGTCAGGGTAAGCCCTTGTGGATCCTTTTGTACCCCTTTTGGAGGTGTCCTCAGAGAAG 778
 QY 369 MetProGlnAlaGlyGlySerMetArgSerThrLeuPheProAsnPheGlySerMetPhe 388
 Db 779 ATGCCTGAGATAAATGGGAAGCATGCGGAGCACATTGTTTCTTAACCTTTGGCAGCATGTTT 838
 QY 389 SerValThrAspGlnHisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAsp 408
 Db 839 AGTGTCGGGAACAGCAGCAAGCTAAAGGTGATTGGGATGCTGAGAGT---CAACGGGAG 895
 QY 409 AspGluGluTyrAlaSerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerPro 428
 Db 896 GGTAAGATTATGGATCAGACCATGGTGGGATGACATTGAAGATAC-CTCCAAGGCCA 954
 QY 429 LeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIle---ValHisHisGlyHis 447
 Db 955 CTTATTCTCGTCAAGCGACAGCGTGGAGGAAGAAAGAGATCGCTGCACCTCATGGCAGT 1014
 QY 448 ArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeuGlyGluGlyGlyAspGlyVal 467
 Db 1015 ATAATGGGTGCTGTGGGA-----AGAAGTAGTAGTCTCATGCAGGGCGGGAGGCAGTA 1068
 QY 468 SerSerThrAspIleGlyGlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGly 487
 Db 1069 AGCAGCATGGGCATTGGTGGGGATGGCAGTTGGCTTGGAAATGGACTGAGAGAGAAGGT 1128
 QY 488 GluAsnGlyArgLysGlyGlyPheLysArgValTyrLeuHisGlnGluGlyValPro 507
 Db 1129 GCAGATGGCGAAAAGAGGTGGCTTCCAACGATCTACTTGCATGAAGAGGGTGTGACA 1198
 QY 508 GlySerArgArgGlySerIleValSerLeuProGlyGlyGlyAspValPheGluGlySer 527
 Db 1189 GGTGATCGCAGGGGCTCTACTACTGTCTATGCTTGCCT---GGAGGTGATGTTCTCTCTGGTGGT 1245
 QY 528 GluPheValHisAlaAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAla 547
 Db 1246 GAGTTCGTCCAGGCAGCTGCTCTTGTGACGCCAACCTGCTCTTTACTTAAGGAATTGATG 1305
 QY 548 GluProArgMetSerAspAlaAlaMetValHisProSerGluValAlaAlaLysGlySer 567
 Db 1306 GAGCAACGCCTTGTGGCCCTGCTATGGTGATGATCCATCTCAGGCAGTGTCTAAAGGTCCA 1365
 QY 568 ArgTrpLysAspLeuPheGluProGlyValArgArgAlaLeuLeuValGlyValGlyIle 587
 Db 1366 AAATGGGACAGACTTATTCGAACCTGGAGTGAAGCATGCTCTGTTTGTGGCATAGGGATA 1425
 QY 588 GlnIleLeuGlnGlnPheAlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeu 607
 Db 1426 CAATCCTGCACAGTTGTGCTGGCATTATGGAGTTCTGTACTACACTCCACAAATTCCT 1485
 QY 608 GluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSer 627
 Db 1486 GAGCAAGCTGGTGTGGTGTCTTCTTGGCAACATTTGGACTTAGTCTCCTCATCTGCATCT 1545
 QY 628 IleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeu 647
 Db 1546 ATTCTTATTAGCGACTGACAACTTGTGATGCTTCCAGCATGGTATTGCTATGAGG 1605
 QY 648 LeuMetAspLeuSerGlyArgArgPheLeuLeuLeuGlyThrIleProIleLeuIleAla 667
 Db 1606 CTCATGGATATGCTGGAAGAAGGTTTCTTCTCTTGGCAACATCCCTATCTCTGATAGTA 1665
 QY 668 SerLeuValIleLeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeu 687
 Db 1666 GCACTAGCTATCTTGATTCTGGTCAATATTCTGGATGTGGGACCCCTGGTTCAATGCCTCA 1725
 QY 688 LeuSerThrValSerValIleValTyrPheCysCysPheValMetGlyPheGlyProIle 707
 Db 1726 CTGTCCACAGTCAGTGTCTACTCTCTCTCTTCTTGTGATGGGTTTGGGCTTATT 1785
 QY 708 ProAsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIle 727

Db 1786 CCAAAACATTCTCTGTGCAGAGATTTCCTCCGACCAACCGTTCGTGGCATCTGCATAGCCATC 1845
 QY 728 CysAlaPheThrPheTrpIleGlyAspIleIleValThrTyrSerLeuProValMetLeu 747
 Db 1846 TGTGCCCTAACATTCTGGATCGGTGATATCATTTGTACATACACCCCTCCCGTGATGCTC 1905
 QY 748 AsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeuIleSerPhe 767
 Db 1906 AACGCCATTGGACTCGCTGGAGTGTTTTGAATCTACGCAGTGGTCTGCATACTGGCTTTC 1965
 QY 768 ValPheValPheLeuLysValProGluThrLysGlyMetProLeuGluValIleThrGlu 787
 Db 1966 CTGTTTGTCTTCATGAAGTGGCGGAGACAAAGGGCATGCCTCTTTGAAGTCATCACCGAG 2025
 QY 788 PhePheAlaValGlyAlaLysGlnAla 796
 Db 2026 TTCTTCTCTGTGCGAGCAAGCAGGCC 2052

RESULT 13
 AC073166 142114 bp DNA linear PLN 28-AUG-2001
 LOCUS Oryza sativa chromosome 10 BAC OSJNB0064P21 genomic sequence,
 DEFINITION complete sequence.
 AC073166
 VERSION AC073166.7 GI:12039314
 KEYWORDS HTG.
 SOURCE Oryza sativa (japonica cultivar-group)
 ORGANISM Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzeae; Oryza.
 REFERENCE 1 (bases 1 to 142114)
 AUTHORS Buell,C.R., Yuan,Q., Moffat,K.S., Hill,J.N., Burr,P.C., Hsiao,J.,
 Zismann,V., Pal,G., Bowman,C.L., Fujii,C.Y., VanAken,S.E.,
 Bowman,C.L., Craven,B., Utterback,T.R., Khalak,H., Feldblyum,T.V.,
 Quackenbush,J., White,O., Salzberg,S.L. and Fraser,C.M.
 Oryza sativa chromosome 10 BAC OSJNB0064P21 genomic sequence
 Unpublished
 2 (bases 1 to 142114)
 Buell,R.
 Direct Submission
 Submitted (09-JUN-2000) The Institute for Genomic Research, 9712
 Medical Center Dr, Rockville, MD 20850, USA
 TITLE 3 (bases 1 to 142114)
 JOURNAL Buell,R.
 Direct Submission
 Submitted (05-JAN-2001) The Institute for Genomic Research, 9712
 Medical Center Dr, Rockville, MD 20850, USA, rbuell@tigr.org
 REFERENCE 4 (bases 1 to 142114)
 AUTHORS Buell,R.
 TITLE Direct Submission
 JOURNAL Submitted (28-AUG-2001) The Institute for Genomic Research, 9712
 Medical Center Dr, Rockville, MD 20850, USA, rbuell@tigr.org
 COMMENT On Jan 5, 2001 this sequence version replaced gi:11968438.
 Address all correspondence to:rice@tigr.org

BAC clone OSJNB0064P21 is from Oryza sativa chromosome 10
 The orientation of the sequence is from SP6 to T7 end of the BAC
 clone.
 Genes were identified by a combination of several methods: Gene
 prediction programs including Genscan and Genscan+ (Chris Burge,
 http://CCR-081.mit.edu/GENSCAN.html), GenemarkHMM (Mark Borodovsky,
 http://genemark.biology.gatech.edu/Genemark/), and Geneslicer
 (Mihaela Pertea and Steven Salzberg, contact mpertea@tigr.org),
 searches of the complete sequence against a peptide database and
 the plant EST database at TIGR (http://www.tigr.org/tdb/cgi.shtml).
 Annotated genes are named to indicate the level of evidence for
 their annotation. Genes with similarity to other proteins are named
 after the database hits. Genes without significant peptide
 similarity but with EST similarity are named as unknown proteins.
 Genes without protein or EST similarity, that are predicted by more
 than two gene prediction programs over most of their length are

annotated as hypothetical proteins. Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, <http://genome.wustl.edu/eddy/tRNAscan-SE/>). Simple repeats are identified by repeatmasker (Arian Smit, <http://ftp.genome.washington.edu/RM/RepeatMasker.html>).

FEATURES		source
gene	1. .142114	
	/organism="Oryza sativa (japonica cultivar-group)"	
	/mol_type="genomic DNA"	
	/cultivar="Nipponbare"	
	/sub_species="japonica"	
	/db_xref="taxon:39947"	
	/chromosome="10"	
	/map="RZ500"	
	/clone="OSJNBb0064P21.1"	
	complement (3008. .3658)	
mRNA	/gene="OSJNBb0064P21.16"	
	/note="predicted by genscan+"	
CDS	complement (<3008. .>3658)	
	/gene="OSJNBb0064P21.16"	
repeat_region	complement (3008. .3658)	
	/gene="OSJNBb0064P21.16"	
repeat_region	/codon_start=1	
	/product="hypothetical protein"	
repeat_region	/protein_id="AAG46103.1"	
	/db_xref="GI:12039315"	
repeat_region	/translation="MVAGDSQARLLVLLALLRLLLDPAAAAAEPelfRRHSDYRATVP ARGISVDFVWAPFESNLRLLHEDRLAPRTPDVLVLSGLWHMLHVTDAARYGDALA SVVDAAKSLRSLPVPVPHFWLGLPLLVNHNLTDAKKVHMNDTILQAYDLEVEQRG LIQRDGGPFLLLDVGLSRGCGQQCTADGMHYDGDVYDAVLHMLNALVIESQQR1"	
	complement (3574. .3596)	
repeat_region	/rpt_family="GC_rich"	
	complement (4269. .4297)	
repeat_region	/rpt_family=" (GGGA)n"	
	5613. .5633	
gene	/rpt_family=" (GGAA)n"	
	complement (5634. .6005)	
mRNA	/gene="OSJNBb0064P21.1"	
	/note="identical to histone H4 (H4C13) GI:168499 (Zea mays), (Plant Physiol Biochem 25, 235 247 (1987)); EST D23040 from this gene"	
CDS	complement (<5634. .>6005)	
	/gene="OSJNBb0064P21.1"	
repeat_region	complement (5634. .5945)	
	/gene="OSJNBb0064P21.1"	
repeat_region	/codon_start=1	
	/product="histone H4"	
repeat_region	/protein_id="AAG46106.1"	
	/db_xref="GI:12039318"	
repeat_region	/translation="MSGRGKGKGLGKGGAHRKVKLVLDNIQGITKPAIRRLARRGGV KRISGLIYEETRGVLKIFLENVIRDVITYTEHARRKVTAMDVVYALKRQRTLYGFG G"	
	complement (7473. .7546)	
repeat_region	/rpt_family=" (CGG)n"	
	7556. .7630	
repeat_region	/rpt_family=" (GGAGAA)n"	
	complement (8238. .8259)	
repeat_region	/rpt_family="AT_rich"	
	complement (8915. .8963)	
repeat_region	/rpt_family="AT_rich"	
	complement (11055. .11097)	
repeat_region	/rpt_family="AT_rich"	
	complement (11244. .11264)	
gene	/rpt_family="AT_rich"	
	14313. .18336	
mRNA	/gene="OSJNBb0064P21.2"	
	/note="identical to calcium dependent protein kinase GI:435466 (Oryza sativa) (Gene 129 (2), 183 189 (1993)); EST C22538, C22540, AA750452, C99271, C22539 from this gene"	
mRNA	Join(14313. .15004,15338. .15481,15564. .15716,15894. .16009, 16646. .16810,17112. .17333,17903. .18336)	
	/gene="OSJNBb0064P21.2"	

Join(14323. .15004,15338. .15481,15564. .15716,15894. .16009, 16646. .16810,17112. .17333,17903. .18025)

/gene="OSJNBb0064P21.2"

/codon_start=1

/product="calcium-dependent protein kinase"

/protein_id="AAG46110.1"

/db_xref="GI:12039322"

/translation="MGNSCQNGTYGNNYQNSNRNFQDRFPASRYVDGNDTDCYSGSSR ASLAGALRQGLNLKSPVLGYKTFNVRELYTLGRELGCQGFQKTYLCTEISTGCOYACK TILKSNLRQVSDIEDVRRITQIMHLSGQKNIVITKDYEQAQVHIVMELCAGGELF SKIQRGHYSERKAAELIKIIVGIIETCHSHGVHRDLKPFENLLLDADDEFSVKALD FGLSVFFRPGQVREVVGSPYIAPEVLEKRYGPEADIWTAGVILYVLLTGTGPPFWAD TQSGIYEKVLDDGRIDFKSNRWPRI SDSAKDLIKMLCPYPSERLKAHEVLKHPWICDN G VATNRALDPSVLPRLKQFSAMNRLKLSLQIIAERLSEEEIVGLREMFKAMDTKNRS VVTFGELKGLKRYSSVFKDTEINDLMEADDTTSTINWEEFIAAAVSLNKIEREKHLM AATYFDKDGSGFITVDKIQKACMERNMEDTFLEEMILEVDQNDGQIDYAEFVTMMQ SNNFGLGWQTVESSLNVALREAPQVY"

complement (18675. .18732)

/rpt_family="GC_rich"

complement (20605. .20634)

/rpt_family="AT_rich"

22393. .24891

/gene="OSJNBb0064P21.6"

/note="predicted by genemark hmm"

<22393. .>24891

/gene="OSJNBb0064P21.6"

22393. .24891

/gene="OSJNBb0064P21.6"

/codon_start=1

/product="hypothetical protein"

/protein_id="AAG46113.1"

/db_xref="GI:12039325"

/translation="MAGSGGEAMASSGLDLGEYLRDPDAIHRRAASVAIVRSGGGDGP RIVDGRDDRRARRSRRLSSWRRPAGGETSVESRRYGFTAAAAAEVGGKTRAAAIW EWKPVRLSRIGKRRCGCLLSVEVAGVRGVPASMDGRLAVTVRKAETKDGAVQTMPA TVRGGCADFVETLFRVCNIYFAGGAGTGKPLKLEPRFRVSVVPAEAPGVRLGAHVD VSSLVDSLQKSSSEGRVRWFDTAVTLSGKATGGELLKLGFLQMLDDAGLCLYTQAAE EKVDVDDVSPALARAHNKNSFSVARTSGPKLSASDAIISPMRAYKQLIDRLSVDEH GDPVTSLIPRKLADDELSGVCLPEYEVVDKQCLDADEEMVTVFVRKLEVVDDKGRKLKQPMT IEAIEALMTNGGKSPSPKVDQKQCLDADEEMVTVFVRKLEVVDDKGRKLKQPMT PRSEKKAATAAPVVPDLGPGLTAVQTRDGGFLVSNPFDLPLEKGDAPPKLMQ VSRPFVLPSSMAATGFDVLQKMAAGGADEVDRKVARLGMDNLTGKTPEQVGFEGIA EAVICGRRTGEGASSAARSVLRKLAAAVSDGRSERVATGIWTAADDPETLEEVI AF SLQKLEAMAVDALMIOAEMADEDAPEFVAPAGDATTVPDALVSPDEWSESRGSDGRV TVVAAIQLRDPSRRYEA VGPVAVVQ SARLLGAAGNSGGRFKVRSLHVGGVQLRCPV GGGAGGGRASWGAERQKLTAMQALAHGPARAAGRRARTPTTTTTPSSQARQRPDVVM SLSSRVLAGMWLKTVRNPDPVRVCATAAGSGGGGN"

complement (23766. .23788)

/rpt_family="GC_rich"

complement (24616. .24639)

/rpt_family="GC_rich"

complement (24693. .24714)

/rpt_family="GC_rich"

24853. .24903

/rpt_family=" (CGG)n"

complement (26149. .26194)

/rpt_family="AT_rich"

complement (26861. .26905)

/rpt_family="AT_rich"

complement (27502. .27537)

/rpt_family="GC_rich"

complement (27628. .27663)

/rpt_family="AT_rich"

complement (28100. .28132)

/rpt_family="AT_rich"

complement (29892. .29929)

/rpt_family=" (TA)n"

29893. .29928

/note="with dinucleotide (AT) repeat (35 to 42 copies)"

34449. .38345

/gene="OSJNBb0064P21.3"

/note="similar to sugar transporter GI:1495273 (Arabidopsis thaliana); EST AA231788, D48171, C72446,

AA749640 from this gene"
Join(34449..34603,34798..34894,35286..35852,35977..36929,
mRNA

Alignment Scores:

Pred. No.: 9.98e-156 Length: 142114
Score: 2514.00 Matches: 552
Percent Similarity: 53.77% Conservative: 75
Best Local Similarity: 47.34% Mismatches: 109
Query Match: 61.69% Indels: 431
DB: 8 Gaps: 9

US-10-051-909-32 (1-800) x AC073166 (1-142114)

Qy	56	MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrp	75
Db	34816	ATGCGGGCGCGTGTGGTCCCATCGCGCTCCATCGGCAACTTGTGCGGGCTGG	34875
Qy	76	AspAsnAlaThrIleAla	81
Db	34876	GATAATGCAACCATTCAGGTAAGTTCGTGGTTAGAGAAAAACGAATCTTTTCT	34935
Qy	81		81
Db	34936	GCCATAGTTTGTGTGCCATGTCGTTATGTGACGCTAGGCGCCGTGTTTGGTGGCAATTAT	34995
Qy	81		81
Db	34996	TGTGATACTTGTGGAGATCATGATGGATTTGGTGGAAATCATCTGCAGTAAAAATAGTTT	35055
Qy	81		81
Db	35056	AGTTGTTCTTCTCTGTTTCATGATATCACAAAAAAACAACTGATATTTTTCAGTA	35115
Qy	81		81
Db	35116	GAATTTGAGTCAATTATGAGTAGGATTAAACAAATGTACGGATCTCATGAATGATCTGT	35175
Qy	81		81
Db	35176	ATTTATGGAACCTTGATACTGCTGCGATCGTAAGGTTTCAGAAATTTGATTCATACCGATG	35235
Qy	82		AlaAlaValle 85
Db	35236	GTTTTGACTTCTTGTGACCCCTGTGGACTTGTGGGATGATCTTTATCAGGTGCGGTACT	35295
Qy	85	uTyrIleLysLysGluPheGlnLeuGlnAsnGluProThrValGluGlyLeuIleValse	105
Db	35296	GTCATCAAGAAGGAATTCAACTTGCAGACGCGCCCTTATCGAAGGCGTGCCTGGC	35355
Qy	105	rMetSerLeuIleGlyAlaThrIleValThrPheSerGlyProLeuSerAspSeril	125
Db	35356	CATGTCGCTCATTTGGGGGACGATCATCACGACGTTCTCTGGAGCAGTGGCTGATCTTT	35415
Qy	125	eGlyArgArgProMetLeuIleLeuSerSerIleLeuTyrPhePheSerGlyLeuIleMe	145
Db	35416	TGCTAGCGGGCCCATGCTGATCGCGTGGCTGCTCTACTTTTGTAGTGGCTAGTAT	35475
Qy	145	tLeuTrpSerProAsnValTyrValLeuLeuAlaArgPheValAspGlyPheGlyil	165
Db	35476	GCTTTGGCGCCAAATGTTGATGTGTGCTCTTGGCGAGGCTCATTTGACGGGTTCGGGAT	35535
Qy	165	eGlyLeuAlaValThrLeuValProLeuTyrIleSerGluIleAlaProSerGluIleAr	185
Db	35536	CGGTTGGCTGTACCGCTTGATACCATTTGTACATCTCTGAGACTGCCCCGACGACATCAG	35595
Qy	185	gGlyLeuLeuAsnThrLeuProGlnPheSerGlySerGlyGlyMetPheLeuSerTyrCy	205
Db	35596	AGGACTGCTAAACACGCTGCCGAGTTTCAGTGGTCTGGAGGGATGTTCTTTCATCTG	35655
Qy	205	sMetValPheGlyMetSerLeuSerProSerProAspTrpArgIleMetLeuGlyValle	225
Db	35656	CATGGTATTTGGCATGTCCTCATGCCACAGCCAGATTGGAGGATCATGCTTGGCGTTCT	35715

Qy	225	uAlaIleProSerLeuPhePhePheGlyLeuThrIlePheTyrLeuProGluSerProAr	245
Db	35716	ATCAATACCATCACTTATATACCTTGCATTGACCATCTTTTACTTACCTGATCGCCGAG	35775
Qy	245	gTrpLeuValSerLysGlyArgMetAlaGluAlaLysLysValLeuGlnLysLeuArgG1	265
Db	35776	GTGGCTCGTGAGCAAGGAAGAATGGCTGAGGCCAAGCGTGTGTGCAAGGCTGCGTGG	35835
Qy	265	yLysAspAspValSer	270
Db	35836	AAGAGAAGATGTTTTCAGGTTGTATGTCTTACCCACCATTTTCCCTGTGAGGATCTAGA	35895
Qy	270		270
Db	35896	AAGTGGAGAAACTATTTAATGCTATCAGTATAGGTATATATCTATCGATGGAGTTTCT	35955
Qy	271		GlyGluLeuSerLeuLeuGluGlyLeuGluValGlyG 284
Db	35956	CATGTTGAATTAACCTGTGCAGGAGAAATGGCCCTTCTCGTTGAAGTCTGGGGTGGGA	36015
Qy	284	lyAspThrSerIleGluGluTyrIleIleGlyProAlaThrGluAlaAlaAspAspLeuV	304
Db	36016	AAGACACAAAATTTAGGAATACATAATTGGACCTGATGAGCTTGTCTGATGAAGGC	36075
Qy	304	alThrAspGlyAspLysGluGlnIleThrLeuTyrGlyProGluGluGlyGlnSerTrpI	324
Db	36076	TGGCTCCA---GATCCAGAGAAGATCAAACTGTATGTCCTGAAGAAGGCTTATCGTGG	36132
Qy	324	leAlaArgProSerLysGlyProIleMetLeuGlySerValLeuSerLeuAlaSerArgH	344
Db	36133	TTGCCCGTCTCTGTTACGGGCAAAAGTGCACCTTGGAAAGTGCAATTAGTCTCATCTCTCGTC	36192
Qy	344	isGlySerMetValAsnGlnSerValProLeuMetAspProIleValThrLeuPheGlyS	364
Db	36193	ATGGTAGTATGTCAGTCAGGGTAAGCCCTTGTGGATCCTGTTGTACCCCTTTTGGAA	36252
Qy	364	erValHisGluAsnMetProGlnAlaGlyGlySerMetArgSerThrLeuPheProAsnp	384
Db	36253	GTGTCCATGAGAAGATGCTGTGAGATAATGGGAAGCATGCGGAGCACATTTCTTAAC	36312
Qy	384	heGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTrpAspGluGluA	404
Db	36313	TTGGCAGCATGTTTAGTGTGGCGGAACAGCAGCAAGCTAAAGGTGATTGGGATGCTGAGA	36372
Qy	404	snLeuHisArgAspAspGluGluTyrAlaSerAspGlyAlaGlyGlyAspTyrGluAspA	424
Db	36373	GT---CAACGGGAGGSGTGAAGATTATGGATCAGACCATGGTGGGATGACATTGAAGATA	36429
Qy	424	snLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIle---V	443
Db	36430	GCCTCCAAAGCCCACTTATTTCTCGTCAAGCCACAAGCGTGGAGGAAGAGATCGCTG	36489
Qy	443	alHisGlyHisArgGlySerAlaLeuSerMetArgGlnThrLeuLeuGlyGluG	463
Db	36490	CACCTCATGGCAGTATAATGGGTGCTGTGGGA-----AGAAGTAGTAGTCTCATGCAGG	36543
Qy	463	lyGlyAspGlyValSerSerThrAspIleGlyGlyTyrGlnLeuAlaTrpLysTrpS	483
Db	36544	CGGGGAGGCAGTAAGCAGCATGGGCATTGTTGGGGGATGGCAGTTGGCTTGGAAATGGA	36603
Qy	483	erGluLysGluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyrLeuHisG	503
Db	36604	CTGAGAGAGAAGGTGCAGATGGCGAAAAAGAGGTGGCTTCCAACGATCTACTTGCATG	36663
Qy	503	lnGluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGlyGlyAspV	523
Db	36664	AAGAGGGTGTGACAGGTATCGCAGGGGCTCTATCTGTCATTGTCCT---GGAGGTGATG	36720
Qy	523	alPheGluGlySerGluPheValHisAlaAlaLeuValSerGlnSerAlaLeuPheS	543
Db	36721	TTCTCTCTGGTGGTGGTTCGTCCAGGCAGTCTCTTGTGACCCAACTGCTCTTTACT	36780
Qy	543	erLysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisProSerGluVala	563

misc_feature

db_xref="taxon:39947"
/chromosome="10"
<1. .12488
/note="Chromosome Sequence Derivation: nucleotide sequence
in this region was derived from BAC clone OSJNBa0040D23
(GB:AC074196)."
<3310. .>4569
/locus_tag="OSJNBa0040D23.23"
/note="similar to chloroplast nucleoid DNA binding protein
CND41 GB:BAA22813 GI:2541876 (Nicotiana tabacum)"
<3310. .>4569
/locus_tag="OSJNBa0040D23.23"
3310. .4569
/locus_tag="OSJNBa0040D23.23"
/codon_start=1
/product="putative nucleoid DNA binding protein"
/protein_id="AAP54833.1"
/db_xref="GI:31433304"
/translation="MGRLLAPMQLLVLCISVTTCAAAGHLRRGLDQGMGRILADAT
AAPGGAVVPLHWSGAHYVANFTTIGTPPQAVSGIVDLSELVWTQCAACRSGCFKQE
LPVFDPSASNTYRAEQCSPLCKSIPTNCSGDGEGYEAPSMFGDTFGIASTDAIAI
GNAEFLGAGCVVASDGSIDGAMDGSGFVGLRTPWSLVGQSNVTAFSCLALHGP
KKSALFLGASAKLAGAKSNPPTPLGQHASNTSDGSDPYTVQLEGIKAGDVAVAA
ASSGGGAIIVLQLETFRLPSYLPDAAYQALEKVVTAAALGSPSMANPEPFDLCFQNA
VSGVPDLVTFQGGATLTAQPSKYLGDGNGNGTVCLSIILSTRLDSADGVSILGSL
LQENVHFLDLEKETLSFEPADCSSLT"
<5409. .>5992
/locus_tag="OSJNBa0040D23.24"
/note="predicted by fgenesh"
join(<5409. .5547,5811. .>5992)
/locus_tag="OSJNBa0040D23.24"
join(5409. .5547,5811. .5992)
/locus_tag="OSJNBa0040D23.24"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAP54834.1"
/db_xref="GI:31433305"
/translation="MTGNLNSNIRICLTPTCILKHPCTGAWMLRQRTSTKAASHCP
TRGKRASKSDLSISAVLDNQNLNHLHGQTKAPLSTIIIVARGSTFVARRNFPFKTKRKN
MYAL"
<8062. .>9183
/locus_tag="OSJNBa0040D23.25"
/note="similar to chloroplast nucleoid DNA binding protein
CND41 GB:BAA22813 GI:2541876 (Nicotiana tabacum)"
<8062. .>9183
/locus_tag="OSJNBa0040D23.25"
8062. .9183
/locus_tag="OSJNBa0040D23.25"
/codon_start=1
/product="putative nucleoid DNA binding protein"
/protein_id="AAP54835.1"
/db_xref="GI:31433306"
/translation="MRGRLLAADAAVAGGGGVPIYWSQPLYMANLTIGTPPQPASAI
IHLAGEFVWTCSPCRRCFKQDLPLFNRSASSTYRPEPCGTALCESVPASTCSGDGVC
SYEVEFMFGDTSGIGTDTFAIGTATASLAFGCAMDSNIKQLLGASGVVGLRTPWSL
VQGMNATPSYCLAPHGAGKKSALLLGASAKLAGGKSAATPLVNTSDSSDYNIHL
EGIKFGDVIAPPPNGSVLVDTIFGVSFLVDAAFQAIKKAVTAVAGAAPMATPTKPF
DLCPFKAAGAGANSLLPLPDVVLTFQGAALTVPPSKYMDAGNGTVCLAMSSAML
NLTTLSILGRHQENIHFLFDLKDXTLSFEPADCSSLS"
<11356. .>12573
/locus_tag="OSJNBa0025B05.4"
/note="similar to chloroplast nucleoid DNA binding protein
CND41 GB:BAA22813 GI:2541876 (Nicotiana tabacum)"
<11356. .>12573
/locus_tag="OSJNBa0025B05.4"
11356. .12573
/locus_tag="OSJNBa0025B05.4"
/codon_start=1
/product="putative chloroplast nucleoid DNA binding
protein"
/protein_id="AAP54836.1"
/db_xref="GI:31433307"
/translation="MGRPVALTLVLCFISVTACSLSEQATGRLLAGVDATPPAAGGA

misc_feature

12489. .23141
/note="Chromosome Sequence Derivation: nucleotide sequence
in this region was derived from BAC clone OSJNBa0025B05
(GB:AC096782)."
complement(<16252. .>17307)
/locus_tag="OSJNBa0025B05.3"
/note="predicted by fgenesh"
complement(<16252. .>17307)
/locus_tag="OSJNBa0025B05.3"
complement(16252. .17307)
/locus_tag="OSJNBa0025B05.3"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAP54837.1"
/db_xref="GI:31433308"
/translation="MVSAAAQAGVVAACVVLVFPVPMGLAGHLSRNKLVFFSGALFVSL
AVGHLSPYLPSLPHLLLAASFHPISSSSSSAASSCVPLLHRVSWADAGGESGV
GRAWAPPSLASTCGLARLRDDASILLNGSWVMVAGDSQARLLVALLRLLDPA
AAAELEFRHRSDYRATVPARGISVDVFWAPFESNLRLHEDRLRLAPRTPDVVLGS
GLWHMLHTDAARYGDALASVVDAAKSLRSLPLVPPPHMFWLGLPLLVNMLNTDAKK
VHMNDTIQAYDLEVEQRLQRDGPFLLLDVGLSRGCGQCTADGMHYDGDVYDA
VLHIMLNALVIESQORI"
complement(<18878. .>19249)
/locus_tag="OSJNBb0064P21.1"
/note="similar to histone H4 (H4C13) GI:168499 (Zea mays),
(Plant Physiol Biochem 25, 235 247 (1987)); EST D23040
from this gene"
complement(<18878. .>19249)
/locus_tag="OSJNBb0064P21.1"
complement(18878. .19189)
/locus_tag="OSJNBb0064P21.1"
/codon_start=1
/product="histone H4"
/protein_id="AAP54838.1"
/db_xref="GI:31433309"
/translation="MSGRGKGGKGLGKGAHRKRVLRNIIQITKPAIRRLARRGGV
KRISGLIYEETRGVLKIFLENVIRDAVITYTEHARRKTVTAMDVVVALKRQGRITLYGFG
G"
<19395. .>20598
/locus_tag="OSJNBa0025B05.2"
/note="similar to B regulatory subunit of PP2A
GB:AAB58900 GI:2160690 (Arabidopsis thaliana)"
join(<19395. .19473,19781. .19910,19992. .>20598)
/locus_tag="OSJNBa0025B05.2"
join(19395. .19473,19781. .19910,19992. .20598)
/locus_tag="OSJNBa0025B05.2"
/codon_start=1
/product="putative protein phosphatase 2A subunit"
/protein_id="AAP54839.1"
/db_xref="GI:31433310"
/translation="MTWLVTTCRRDAIRGRVLRHWRMDGAGPPGDDGAGQDGRRLQ
VPASGQCRPPTRRRRPTCTSTRTLRNHVDRSFLAGLVALFAFEDSRERDRLKIV
YHQLYSLTVERAFMRIMAAALLRFVYDTSPEAEHRCGVGELLEICGSIINGFAVP
LKEEHRAPMARVLLPLHRTRWHTYRQLSYCGSIRRSRHLPPPLAGDQMPGGAHR
RARGDRDPLSAPGLPAASAAAHRFVCFIHFHNRHFHSQKKIDFIYICTFN"
23142. .155358
/note="Chromosome Sequence Derivation: nucleotide sequence
in this region was derived from BAC clone OSJNBb0064P21
(GB:AC073166)."
<27557. .>31580
/locus_tag="OSJNBb0064P21.2"
/note="similar to calcium dependent protein kinase
GI:435466 (Oryza sativa) (Gene 129 (2), 183 189 (1993));
EST C22538, C22540, AA750452, C99271, C22539 from this
gene"
join(27557. .28248,28592. .28725,28808. .28960,29138. .29253,

CDS
29890. 30054,30356..30577,31147..31580)
/locus_tag="OSJNB0064P21.2"
join(27567..28248,28582..28725,28808..28960,29138..29253,
29890..30054,30356..30577,31147..31269)
/locus_tag="OSJNB0064P21.2"
/codon_start=1
/product="calcium-dependent protein kinase"
/protein_id="AAP54840.1"
/db_xref="GI:31433311"
/translation="MGNSCQNGTYGNNYQNSNRFQNDRFASRYVDGNDTDCYSGSSR
ASLAGALRQGLNLKSPVLGYKTPNVRELYTLGRELQCGQFGKTYLCTEISTGCGYACK
TILKSNLRCSVSDIEDVRREIQIMHLSGQKNIVTKDYEDEQAVHIVNELCAGGELF
SKIQRGHYSEKAAELIKIIVGIIETCHSHGVNHRDLKPENFLLLDADDEFVSKAID

Alignment Scores:		Length: 300957	
Pred. No.:	2,96e-155	Matches:	552
Score:	2514.00	Conservative:	75
Percent Similarity:	53.77%	Mismatches:	109
Best local Similarity:	47.34%	Indels:	431
Query Match:	61.69%	Gaps:	9
DB:	8		
US-10-051-909-32 (1-800) x AE017116 (1-300957)			
QY	56	MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrp	75
Db	48060	ATGGCGGCGCGCTGCTGGTGGCCATCGCGGCTCCATCGGCAACTTGCTGCAGGGCTGG	48119
QY	76	AspAsnAlaThrIleAla-----	81
Db	48120	GATAATGCAACCATTCAGGTAAGTTCGTGGTTAGAGAAAAACCAATCTTTTCT	48179
QY	81	-----	81
Db	48180	GCCATAGTTTGTGTGCCATGTCTGTTATGTGACGCTAGGGCCCGTGTGGTGGCAATTAT	48239
QY	81	-----	81
Db	48240	TGTGATACTTGTGAGATCATGATGGATTGGTGAATCATCTGCAGTAAAAATAGTTTT	48299
QY	81	-----	81
Db	48300	AGTTGTTCTTCTGTTCATGATATCACAAAAAAAACAACCTGATATTTTTTTCAGTA	48359
QY	81	-----	81
Db	48360	GAATTTTGAGTCAATTATGAGTAGGATTAAACAAATGTACGGATCTCATGAATGATCTGT	48419
QY	81	-----	81
Db	48420	ATTTATGGAACITGATACTGTCTGCATCGTAAGGTTTCAGAAATTTGATTGATACCGATG	48479
QY	82	-----	85
Db	48480	GTTTTGACTTCTTGCTGACCCCTGTGGACTTGTGGGATGATCTTTATCAGGTGCGGTACT	48539
QY	85	uTyriLeLysGluPheGlnLeuGlnAsnGluProThrValGluGlyLeuIleValse	105
Db	48540	GTACATCAAGAAAGGAATTCAACTTGCAGACGAGCCCTTATCGAAGGCCTGATCGTGGC	48599
QY	105	rMetSerLeuIleGlyAlaThrIleValThrPheSerGlyProLeuSerAspSerI1	125
Db	48600	CATGTCGCTCAITGGGGCGACGATCATCACACGTTCTCTGGAGCAGTGGCTGATCTTT	48659
QY	125	eGlyArgProMetLeuIleLeuSerSerIleLeuTyrrPhePheSerGlyLeuIleMe	145
Db	48660	TGGTAGCGGGCCCATGCTGATCGCTCGGCTGTCTCTACTTTTGTAGTGGCTAGTGAT	48719
QY	145	tLeuTrpSerProAsnValTyrrValLeuLeuAlaArgPheValAspGlyPheGlyI1	165
Db	48720	GCTTTGGCGCCAAATGTGTATGTGTGCTCTTGGCGAGGCTCATTGACGGTTTCGGGAT	48779
QY	165	eGlyLeuAlaValThrLeuValProLeuTyrrIleSerGluIleAlaProSerGluIleAr	185

Db	48780	CGGTTTGGCTGTACGCTTGTACCATTTGTACATCTCTGAGACTGCCCCGACGACATCAG	48839
QY	185	gGlyLeuLeuAsnThrLeuProGlnPheSerGlySerGlyGlyMetPheLeuSerTyrrCy	205
Db	48840	AGGACTGCTAAACACGCTGCCGCAGTTCACTGGGTCTGGAGGGATGTCTTTCATACTG	48899
QY	205	sMetValPheGlyMetSerLeuSerProSerProAspTrpArgIleMetLeuGlyValle	225
Db	48900	CATGGTATTGGCATGTCCTCATGCCACAGCCAGATTGGAGGATCATGCTTGGCGTTCT	48959
QY	225	uAlaIleProSerLeuPhePhePheGlyLeuThrIlePheTyrrLeuProGluSerProAr	245
Db	48960	ATCAATACCATCACATTATATCTTTTGCATTGACCATCTTTTACCTGAATCGCCGAG	49019
QY	245	gTrpLeuValSerLysGlyArgMetAlaGluAlaLysLysValLeuGlnLysLeuArgG1	265
Db	49020	GTGGCTCGTGAGCAAGGAAGAAATGGCTGAGGCCAAGCGTGTGTGCAAGGCCTGCGTGG	49079
QY	265	yLysAspAspValSer-----	270
Db	49080	AAGAGAAGATGTTTCAGGTTTGTATGTCTTACCCAGCATTTTTTCCCTGTGAGGATCTAGA	49139
QY	270	-----	270
Db	49140	AAGTGGAGAAACTATTAAATGCTATCAGTATAGGTATATTATCTATCGCATGGAGTTTCT	49199
QY	271	-----GlyGluLeuSerLeuLeuGluGlyLeuValGlyG	284
Db	49200	CATGTTGAATTAACGTGTGCAGGAGAAATGGCCCTTCTCGTTGAAGGTCTGGGGTTGGGA	49259
QY	284	lyAspThrSerIleGluGluTyrrIleGlyProAlaThrGluAlaAlaAspAspLeuV	304
Db	49260	AAGACACAAAAATTGAGGAATACATAATTGGACCTGATGATGAGCTTGCTGATGAAGGC	49319
QY	304	alThrAspGlyAspLysGluGlnIleThrLeuTyrrGlyProGluGluGlnSerTrpI	324
Db	49320	TGGCTCCA--GATCCAGAGAAGATCAAACTGTATGTTCTCTGAAGAGGCTTATCGTGG	49376
QY	324	leAlaArgProSerLysGlyProIleMetLeuGlySerValLeuSerLeuAlaSerArgH	344
Db	49377	TTGCCCGTCTCTGTTCAAGGCAAGTGCACITGGAAGTGCAATTAGTCTCATCTCTCGTC	49436
QY	344	isGlySerMetValAsnGlnSerValProIleMetAspProIleValThrLeuPheGlys	364
Db	49437	ATGGTAGTATGTTCACTCAGGGTAAGCCCTTGTGGATCCTGTGTGTCACCCCTTTTGGAA	49496
QY	364	erValHisGluAsnMetProGlnAlaGlyGlySerMetArgSerThrLeuPheProAsnP	384
Db	49497	GTGTCATGAGAAAGATGCTTGAGATAATGGGAAGCATGCGGAGCACATTGTTCTTAAC	49556
QY	384	heGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTrpAspGluGluA	404
Db	49557	TTGGCAGCATGTTTGTGTCGGGAACAGCAGCAAGCTAAAGGTGATTGGGATGCTGAGA	49616
QY	404	snLeuHisArgAspAspGluGluTyrrAlaSerAspGlyAlaGlyGlyAspTyrrGluAspA	424
Db	49617	GT---CAACGGGAGGGTGAAGATTATGGATCAGACCATGTTGGGGATGACATTGAAGATA	49673
QY	424	snLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIle---	443
Db	49674	GCCTCAAAGCCCACTTATTCTCGTCAAGCGACAAGCGTGGGAAGGAAGGAGATCGCTG	49733
QY	443	alHisHisGlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeuGlyGluG	463
Db	49734	CACCTCATGGCAGTATAATGGTGTGTGGGA-----AGAAGTAGTAGTCTCATGCAGG	49787
QY	463	lyGlyAspGlyValSerSerThrAspIleGlyGlyTrpGlnLeuAlaTrpLysTrps	483
Db	49788	GCGGGAGGCAGTAAGCAGCATGGGCATTGTTGGGGGATGGCAGTTGGCTTGGAAATGGA	49847
QY	483	erGluLysGluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyrrLeuHisG	503

/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/variety="Columbia"
/db_xref="taxon:3702"
/chromosome="4"
173. .1890
/gene="F23E12.10"
join(<173. .374,456. .919,1089. .1610,1678. .1890)
/gene="F23E12.10"
/codon_start=1
/product="hypothetical protein (fragment)"
/protein_id="CAA18726.1"
/db_xref="GI:3080407"
/db_xref="GOA:O65486"
/db_xref="TrEMBL:O65486"
/translation="FNAAVAGCFFAVESVLWPSSTDSSTSLPNTTSMVILSAVTASV
VSEIGLSEPAFKVPDYDFRSPGELPLYLLGALCGLVSLASRCTSSMTSAVDSLNK
DAGIPKAVFPVMGGLSVGIIALVYPEVLWGFQNVVDILLEKRPFKGLSADLLQLVA
VKIAATAWCRASGLVGYAPSLFIGGAAGMAYGKFIGLALAQNPFLNLSILEVASPQ
AYGLVGMATLAGVCQVPLTAVLLLFELTQDYRIVLELLGAVGMSWITSGQSKROET
RETKETRKRSQEAQVSLTSSDDSESTNNLCEVESLCLDDSLNQSEELPKSIFVSEA
MRTRFATVMSTSLAEALTRMLIEKQSCALIVDPDNIFLGILTLSDIQEFKARKEGN
NRPKIDIFVNDICSRSGGCKVPWTVPDMDLLAAQTIMNKHELSHVAVVSGSIDAPRI
HPVGVDRECIITLR"
173. .374
/gene="F23E12.10"
/number=1
375. .455
/gene="F23E12.10"
/number=1
456. .919
/gene="F23E12.10"
/number=2
920. .1088
/gene="F23E12.10"
/number=2
1089. .1610
/gene="F23E12.10"
/number=3
1611. .1677
/gene="F23E12.10"
/number=3
1678. .1890
/gene="F23E12.10"
/number=4
3124. .4428
/gene="F23E12.20"
join(3124. .3241,3323. .3465,3595. .3696,3945. .4107,
4211. .4428)
/gene="F23E12.20"
/note="similarity to dihydroflavonol-4-reductase, Gerbera
x sp., PIR2:S35189"
/codon_start=1
/product="putative protein"
/protein_id="CAA18727.1"
/db_xref="GI:3080408"
/db_xref="TrEMBL:O65487"
/translation="MDQAKGKVCVTGASGFLASWLVKRLLEGYEVIGTVRDPGNEKK
LAHLWKLEGAKERLRVLVKADLMEEGSDNAIMGQGVFHTASPEILRPAIEGTLNVL
SCRKNPSLKRVLTTSSSTVWYALSKTLAEQAARKFSENGIDLTVLPFLVGPSP
PDLGSTASDVLGLLKGETEKFGWQMGYVHIDDVARHTIVVFEHEAAQGRYICSSNV
ISLEELVSFLSARYPSLPIPKRFDTDPLN"
3124. .3241
/gene="F23E12.20"
/number=1
3242. .3322
/gene="F23E12.20"
/number=1
3323. .3465
/gene="F23E12.20"
/number=2
3466. .3594
/gene="F23E12.20"

/number=2
3595. .3696
/gene="F23E12.20"
/number=3
3697. .3944
/gene="F23E12.20"
/number=3
3945. .4107
/gene="F23E12.20"
/number=4
4108. .4210
/gene="F23E12.20"
/number=4
4211. .4428
/gene="F23E12.20"
/number=5
4939. .6163
/gene="F23E12.30"
complement(join(4939. .5013,5109. .5195,5307. .5342,
5451. .5651,5725. .5811,6161. .6163))
/gene="F23E12.30"
complement(join(4939. .5013,5109. .5195,5307. .5342,
5451. .5651,5725. .5811,6161. .6163))
/gene="F23E12.30"
/note="Contains Clathrin adaptor complexes small chain
signature [VVYKRYASLYF]
contains EST gb:T44219"
/codon_start=1
/product="clathrin assembly protein Ap19 homolog"
/protein_id="CAA18728.1"
/db_xref="GI:3080409"
/db_xref="GOA:O23685"
/db_xref="TrEMBL:O23685"
/translation="MIHFVLLVSRQGVRLTKWYSPYAKERSKVIRELSGVILNRGP
KLCNFVWGRGYKVYKRYASLYFCMCIDQEDNELELEIHHYVEILDYRFGSVCELD
LIFNFKAYYILDELLIAGELQESSKKTVARIIISAQDLVEVAKEEASSISNIIAQAT
NR"
complement(4939. .5013)
/gene="F23E12.30"
/number=1
complement(5014. .5108)
/gene="F23E12.30"
/number=1
complement(5109. .5195)
/gene="F23E12.30"
/number=2
complement(5196. .5306)
/gene="F23E12.30"
/number=2
complement(5307. .5342)
/gene="F23E12.30"
/number=3
complement(5343. .5450)
/gene="F23E12.30"
/number=3
complement(5451. .5651)
/gene="F23E12.30"
/number=4
complement(5652. .5724)
/gene="F23E12.30"
/number=4
complement(5725. .5811)
/gene="F23E12.30"
/number=5
complement(5812. .6160)
/gene="F23E12.30"
/number=5
complement(6161. .6163)
/gene="F23E12.30"
/number=6
7352. .7582
/gene="F23E12.40"
complement(7352. .7582)

```

CDS
    /gene="F23E12.40"
    complement(7352..7582)
    /gene="F23E12.40"
    /codon_start=1
    /product="hypothetical protein"
    /protein_id="CAA18729.1"
    /db_xref="GI:3080410"
    /db_xref="TrEMBL:O65488"
    /translation="MPLVRKNGVVMRIGSPMLHSENLVGDGVEAPPRGERDEIGEEV
GGETNVVSESCGGEQAQDAKTTPEGEADRI"
    complement(7352..7582)
    /gene="F23E12.40"
    /number=1
    8233..9045
    /gene="F23E12.50"
    complement(8233..9045)
    /gene="F23E12.50"
    complement(8233..9045)
    /note="similarity to DNA-binding protein PDI, Pisum
sativum"
    /codon_start=1
    /product="putative protein"
    /protein_id="CAA18730.1"

Alignment Scores:
Pred. No.:      3.09e-155      Length:      86710
Score:          2502.00      Matches:      523
Percent Similarity: 68.91%      Conservative: 91
Best Local Similarity: 58.70%      Mismatches: 124
Query Match:      61.40%      Indels:      153
DB:               8          Gaps:      12

US-10-051-909-32 (1-800) x ATF23E12 (1-86710)
QY 52 LeuGluAspLys---MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsn 70
Db 39649 ATAGAAGACGGAATTATGAGTGGAGCTGTGCTTGTCTATTGCTGCTGCTGTTGGCAAC 39708
QY 71 LeuLeuGlnGlyTrpAspAsnAlaThrIle-----80
Db 39709 TTGTTACAAGGATGGGATAACGCACTATTGCGAGGTGATTGATTATCTTCTGTATCCATA 39768
QY 80 -----80
Db 39769 ATTCTGTTTATGAATTAGGATCGATAAAGTCAGGATTTTATGGATACAATTTCACAGA 39828
QY 81 -----Al 81
Db 39829 TCTTTTAGCTTATCTCTTTTGCTTGCTGATGAGTTTAAAGTTTCTTTGGGTACTCTGC 39888
QY 81 aAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsnGluProThrValGluG1 101
Db 39889 AGGAGCTGTGTTGTACATAAATAAGGAGTTAATTGGAGAGTAAATCCATCAGTGGAAAG 39948
QY 101 yLeuIleValSerMetSerLeuIleGlyAlaThrIleValThrPheSerGlyProLe 121
Db 39949 TCTAATTGGCGGATGCACCTATTATGGTGCTCTCTGATTACAAACATGCTCTGGAGGGGT 40008
QY 121 uSerAspSerIleGlyArgProMetLeuIleLeuSerSerIleLeuTyrPhePheSe 141
Db 40009 AGCTGATTGGCTTGGTCGCCGCTCCCATGCTAAATATTGTCCTCAATCTCTACTTTGTTGG 40068
QY 141 rGlyLeuIleMetLeuTrpSerProAsnValTyrValLeuLeuLeuAlaArgPheValAs 161
Db 40069 TTCTCTAGTAATGCTATGGTCTCCGAATGTTTATGTGTTGCTCTTAGGAAGTTTGTAGA 40128
QY 161 pGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyrIleSerGluIleAlaPr 181
Db 40129 TGGATTTGGGGTGGTCTTGTGGTCACTGTTCTCTATTATATATCTGAGACTGCACC 40188
QY 181 oSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSerGlySerGlyGlyMetPh 201

```

```

Db 40189 ACCTGAGATTAGGGAGCTGTTGAATACGCTACCGCAGTTCACCTGGCTCTGGAGGGATGTT 40248
QY 201 eLeuSerTyrCysMetValPheGlyMetSerLeuSerProSerProAspTrpArgIleMe 221
Db 40249 CTTATCTTACTGTATGTTTTCGGAATGTCGTTGATGCCATCACCTAGCTGGAGATTGAT 40308
QY 221 tLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeuThrIlePheTyrLeuPr 241
Db 40309 GCTTGGTGTCTTTTCATCCCTTCCCTTGTCTTTTCTCCTCACGGTCTTCTTCTTCC 40368
QY 241 oGluSerProArgTrpLeuValSerLysGlyArgMetAlaGluAlaLysLysValLeuG1 261
Db 40369 CGAGTCCCCAAGGTGGCTCGTCGAGCAAGGTCGAATGCTTGAAGCAAAGCGGTTCTTCA 40428
QY 261 nLysLeuArgGlyLysAspValSer-----270
Db 40429 GAGACTGCGTGGTCGCGAGAGATGTGCTGTTAGTCTTAAATCGTTCTGACTCTAATTCT 40488
QY 271 -----GlyGluLeuS 274
Db 40489 TTCAATGAAACCAATCTTTTCTATTGAATGTTCTTATGTGGCAATGTATAGTGAGATGG 40548
QY 274 erLeuLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIleG 294
Db 40549 CTTTGTGTTGAGGGTCTTGGAAATTGGAGGTGAACAACCATAGAGGATATATAATTG 40608
QY 294 lyProAlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThrL 314
Db 40609 GTCCCGCGGATGAAGTTACTGATGATCATGATATAGCTGTGGATAAGGATCAAATTAAAGT 40668
QY 314 euTyrGlyProGluGluGlnSerTrpIleAlaArgProSerLysGlyProIleMetL 334
Db 40669 TATATGGTGCAGAAAGGGCTGAGTTGGTGTCTAGGCCAGTCAAAGGA-----40718
QY 334 euGlySerValLeuSerLeuAlaSerArgHisGlySer---MetValAsnGlnSerValP 353
Db 40719 --GGAAGCACTATGAGTGTGTTTGTCTCGCCATGGAAGTACAATGAGCAGGAGCAAGGCT 40776
QY 353 roLeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaG 373
Db 40777 CATTGATTGATCCTCTGTGCACACTGTTTGGAGCGTTCACGAGAAGATGCCGACACT- 40835
QY 373 lyGlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspG 393
Db 40836 --GGAAGCATGAGAGTGCCTTGTTCACACTTGTCCACATTTTGGAGTATGTTGAGGGA 40893
QY 393 lnHisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspGluGluTyrA 413
Db 40894 ATCAACCAAGACATGAAGATTGGGATGAAGAGAATCTTGTGGAGAAGGTGAGGATTATC 40953
QY 413 laSerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgG 433
Db 40954 CATCCGAC---CATGGAGATGATTCTGAAGATGATCTTCATCTCCGTGATCTCACGTC 41010
QY 433 lnAlaThrGlyAlaGluGlyLysAspIleValHisGlyHisArgGlySerAlaLeuS 453
Db 41011 AAACGACAAGCATGGAG--AAAGACATGCTCACACTGCTCATGGAACCTTTTCTACCT 41067
QY 453 erMetArgArgGlnThrLeuLeuGlyGlyGlyAspGlyValSerSerThrAspIleG 473
Db 41068 TCAGACATGGAAGTCAAGTGCAGGGAGCTCAAGGGGAAGGAGCGGTAGTATGGGGATTG 41127
QY 473 lyGlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArgLysG 493
Db 41128 GAGGTGGATGGCAAGTGGCATGGAAATGGACGGAAGAGAAGATGAATCGGGACAGAAAG 41187
QY 493 luGlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArgGlys 513
Db 41188 AAGGTGGGTTTAAACGGATATATCTTGCATCAAGAAGGTTTCCAGGATCTCGACGTGGCT 41247
QY 513 erIleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaA 533
Db 41248 CAATTGTTTTCATTCCTGCTGGTGTGATGGAACCGGTGAG--GCAGATTTTGTACAAGCGCT 41304

```

QY 533 laAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerA 553
Db :|||||
41305 CTGCTTGGTTAGCCAAACCAGCTCTTTATTTCCAAAGACCTTCTCAAAGAACATACAATTG 41364
QY 553 spAlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTirLysAspLeuP 573
Db :|||||
41365 GTCCTGTATGGTACATCCATCCGAA--ACAACATAAGGTCATTTGGCATGATCTTC 41421
QY 573 heGluProGlyValArgAlaLeuValGlyValGlyIleGlnIleLeuGln--- 591
Db :|||||
41422 ATGATCCTGGAGTCAAGCGTGCATTAGTCGTAGGAGTTGGACTTCAAATACCTTCAGCAGG 41481
QY 591 ----- 591
Db 41482 TAGTCTAGATAGCCACATCAAAATTAGTGAATCACAGAAGAAATAATGCGGTTTATGCTA 41541
QY 592 -----GlnPheAlaGlyIleAsnGlyValLeuT 601
Db 41542 ATTCTATGTTTTCCTGCTTCTCTGGAATGCAGTTCTCTAGGCATCAACGGAGTTCCTT 41601
QY 601 yrTyrThrProGlnIleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyL 621
Db :|||||
41602 ACTACACACCGCAAAATCCTTGAGCAGCGGGGTGTCGGGATCCTACTATCGAACATGGGGA 41661
QY 621 euSerSerAlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProC 641
Db :|||||
41662 TTAGTTCCTCCTCAGCATCCTTACTTATAAGTGCAATTGACAACTTTGTGATGTTACCTG 41721
QY 641 ysileGlyPheAlaMetLeuLeuMetAspLeuSerGlyArg----- 654
Db 41722 CAATAGCTGTGCAATGAGGCTCATGGATCTTTCTGTCGAAGGTGATTAAACTTAAAT 41781
QY 654 ----- 654
Db 41782 TTCCTTTGTTCTTACAATACACATAAACGAAACAAAGCTAAATGCGTTATGAAATGCTT 41841
QY 655 -----ArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeu 669
Db 41842 TTATTTTATATTGTCAGGACCTTGCTTCTCACCACGATACCAATCCTGATAGCATCTCTA 41901
QY 670 ValIleLeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSer 689
Db :|||||
41902 TTGGTTTATAGTAATCTCAAATCTTGTTCACATGAACAGCATTGTGCACGCGGTCTTATCA 41961
QY 690 ThrValSerValIleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsn 709
Db 41962 ACCGTAAGCGTTGCTCTACTTCTGCTTCTCGTGATGGGTTTCGGTCTCTGCCAAAC 42021
QY 710 ileLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAla 729
Db 42022 ATCCTCTGTTCAGAGATTTTCCAACTCGAGTCCGCGGAATCTGCATCGCCATCTGGCA 42081
QY 730 PheThrPheTrpIleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAla 749
Db 42082 CTCACCTTCTGGATCTGTGACATAATCGTCACCTACAGTCTCCCCGCTGCTCAATCC 42141
QY 750 IleGlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPhe 769
Db 42142 ATTGGACTAGCTGGTGTGTTTGAATGTACGCAATCGTATGTTGCAITTCATGGGTCTTT 42201
QY 770 ValPheLeuLysValProGluThrLysGlyMetProLeuGluValIleThrGluPhePhe 789
Db 42202 GTGTTTCAATTAAGTCCCGAAACTAAAGGCATGCCACTTGAAGTCATCACAGAGTTCCTT 42261
QY 790 AlaValGlyAlaLysGlnAlaAla 798
Db 42262 TCTGTTGGAGCTAGACAAAGCTGAAGCT 42288

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 13, 2004, 11:39:19 ; Search time 978.566 Seconds
(without alignments)
4291.521 Million cell updates/sec

Title: US-10-051-909-32
Perfect score: 4075
Sequence: 1 IRSGSWLAVQTPTPLDRR.....PLEVITEFFAVGAKQAAKA 800

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool_h/US10051909/runat_13102004_123335_19900/app_query_fasta_1.1678
-DB=N_Geneseq_23Sep04 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10051909 @CGN 1 1 891 @runat_13102004_123335_19900 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_23Sep04:
1: geneseqn1980s:
2: geneseqn1990s:
3: geneseqn2000s:
4: geneseqn2001as:
5: geneseqn2001bs:
6: geneseqn2002as:
7: geneseqn2002bs:
8: geneseqn2003as:
9: geneseqn2003bs:
10: geneseqn2003cs:
11: geneseqn2003ds:
12: geneseqn2004s:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4075	100.0	2777	12	ADG47935
2	4059	99.6	2908	12	ADM47932
3	3517	86.3	2824	6	ABK51962
4	3517	86.3	2824	8	ABX93198
5	3517	86.3	2824	12	ADG47905
6	2674	65.6	2601	6	ABK51965

7	2674	65.6	2601	8	ABX93201	Abx93201 cDNA enco
8	2674	65.6	2601	12	ADG47911	Adg47911 Corn Arab
9	2584	63.4	2190	6	ABZ13510	Abz13510 Arabidops
10	2275	55.8	2205	6	ABZ14449	Abz14449 Arabidops
11	2275	55.8	2205	6	ADG87978	Adg87978 A. thalia
12	1670.5	41.0	1692	6	ABK51966	Abk51966 Soybean c
13	1670.5	41.0	1692	8	ABX93202	Abx93202 cDNA enco
14	1670.5	41.0	1692	12	ADG47913	Adg47913 Soybean A
15	1424	34.9	1487	6	ABK51968	Abk51968 Wheat con
16	1424	34.9	1487	8	ABX93204	Abx93204 cDNA enco
17	1424	34.9	1487	12	ADG47917	Adg47917 Wheat Ara
18	1035	25.4	1412	12	ADJ39747	Adj39747 Plant CDN
19	943	23.1	1009	6	ABK51969	Abk51969 Wheat CDN
20	943	23.1	1009	8	ABX93205	Abx93205 cDNA enco
21	943	23.1	1009	12	ADG47919	Adg47919 Wheat Ara
22	620	15.2	751	12	ADJ41683	Adj41683 Plant CDN
23	616	15.1	1806	10	ADC07791	Adc07791 Rice DNA
24	600	14.7	778	12	ADJ41684	Adj41684 Plant CDN
25	587.5	14.4	1518	6	ABL41880	Ab141880 Nucleotid
26	562	13.8	870	6	ABK51964	Abk51964 Rice cDNA
27	562	13.8	870	8	ABX93200	Abx93200 cDNA enco
28	562	13.8	870	12	ADG47909	Adg47909 Rice Arab
29	552	13.5	659	12	ADJ41685	Adj41685 Plant CDN
30	548	13.4	644	12	ADJ42193	Adj42193 Plant CDN
31	539	13.2	1395	6	ABK73616	Abk73616 Bacillus
32	534	13.1	2127	12	ADM47936	Adm47936 Polynucle
33	520.5	12.8	1826	3	AAC42332	Aac42332 Arabidops
34	496	12.2	1853	6	ABK51973	Abk51973 Soybean c
35	496	12.2	1853	8	ABX93209	Abx93209 cDNA enco
36	496	12.2	1853	12	ADG47927	Adg47927 Soybean B
37	494.5	12.1	2017	6	ABK51972	Abk51972 Rice cont
38	494.5	12.1	2017	8	ABX93208	Abx93208 cDNA enco
39	494.5	12.1	2017	12	ADG47925	Adg47925 Rice Beta
40	493.5	12.1	1527	3	AAC43261	Aac43261 Arabidops
41	493	12.1	1482	6	ABZ12990	Abz12990 Arabidops
42	493	12.1	1482	12	ADN72404	Adn72404 Thale cre
43	492	12.1	1338	8	ACA47151	ACA47151 Prokaryot
44	491.5	12.1	1914	6	ABK51971	Abk51971 Corn cDNA
45	491.5	12.1	1914	8	ABX93207	Abx93207 cDNA enco

ALIGNMENTS

RESULT 1
ADG47935
ID ADG47935 standard; cDNA; 2777 BP.
AC ADG47935;
DT 11-MAR-2004 (first entry)
XX Corn Arabidopsis-like sugar transport protein cDNA #2.

Arabidopsis-like sugar transport protein;
Beta-vulgaris-like sugar transport protein; transgenic; physical mapping;
corn; plant; gene; ss.

Zea mays.

Key Location/Qualifiers
CDS 12..2414

FT /*tag= a

FT /product= "Corn Arabidopsis-like sugar transport protein"

FT /note= "No start codon"

FT /partial

XX US2002199217-A1.

XX 26-DEC-2002.

XX 17-JAN-2002; 2002US-00051909.

XX 24-APR-1998; 98US-0083044P.

PR 14-APR-1999; 99US-00291922.
XX (HELE/) HELENTJARIS T G.
XX PI Helentjaris TG;
XX WPI; 2004-040967/04.
DR P-PSDB; ADG47936.
XX
PT New isolated polynucleotide encoding a polypeptide having sugar transport
PT protein activity, for producing a transformed plant and for use as probes
PT in physical mapping.
XX
PS Claim 6; SEQ ID NO 31; 71pp; English.
XX
CC The invention relates to Arabidopsis-like or Beta-vulgaris-like sugar
CC transport proteins and their corresponding nucleic acid sequences. The
CC sequences of the invention are useful to transform a cell. These are also
CC useful to produce a transgenic plant. Probes derived from sequences
CC encoding sugar transport protein may be used for physical mapping. The
CC present sequence is corn Arabidopsis-like sugar transport protein cDNA.
XX
SQ Sequence 2777 BP; 667 A; 587 C; 701 G; 822 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 0 Length: 2777
Score: 4075.00 Matches: 800
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0
US-10-051-909-32 (1-800) x ADG47935 (1-2777)
QY 1 IleArgSerGlySerTrpLeuAlaValGlnThrProPheThrProAspLeuAspArgArg 20
Db 12 ATTCGAGCGGCTCTTGGCTTGCAGTCCAAACGCCCTTACCCCTGATCGACCGGAGG 71
QY 21 GluArgLeuLeuProSerValValLeuAlaLeuProGlyProLeuProAlaSerCys 40
Db 72 GAGCGGCTCCTCCGTCAGTGTCTTGTCTTGGCTGGGCTCTTCCGCTGCTTCTGT 131
QY 41 SerSerGlnGluProValThrSerAspAspIleLeuGluAspLysMetSerGlyAlaVal 60
Db 132 TCCTCACAGGAGCGGTCGACCTCGGACGATATCTTGGAGGACAAGATGTCGGGGCTGT 191
QY 61 LeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrpAspAsnAlaThrIle 80
Db 192 CTGTGCGCCATAGTCGCTCCATCGGCAATCTATTGACGGGTGGGACAATGCCACCATC 251
QY 81 AlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsnGluProThrValGlu 100
Db 252 GCAGCTGCTGTTCTGTATATAAAGAGGAATTTCAATTGCAAAATGAGCCCACTGTGGAG 311
QY 101 GlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThrThrPheSerGlyPro 120
Db 312 GGACTAATTGTGTCATGTCACTTATCGGCGCCACCATCGTTACTACTACTTCCGGGCCA 371
QY 121 LeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSerIleLeuTyrPhePhe 140
Db 372 TTATCAGACTCGATTGSCCGACGCCCTATGCTTATCTCTCTTCAATTCTGTACTTCTTC 431
QY 141 SerGlyLeuIleMetLeuTrpSerProAsnValTyrValLeuLeuAlaArgPheVal 160
Db 432 AGCGGCTCATCATGCTATGGTCTCTCTAATGTCTATGCTGCTGTGTGGACGCTTCGTA 491
QY 161 AspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyrIleSerGluIleAla 180
Db 492 GATGGATTGGTATTGGCTTGGCTGTCACGCTTGTGCTTGTGCTTGTACATTTCAGAAATAGCC 551
QY 181 ProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSerGlySerGlyMet 200
Db 552 CCTTCGGAGATTAGAGGTTTGTGTAATACACTACCAATTCAGTGTGATCAGGAGGAATG 611

QY 201 PheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSerProAspTrpArgIle 220
Db 612 TTCTTGTCTACTGTCATGGTGTGGATGTCCTGTGCGCATCACCGATTGGAGATT 671
QY 221 MetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeuThrIlePheTyrLeu 240
Db 672 ATGCTTGGTGTGCTCGGATACCTTCAATGTTCTTCTTGGTTGACAAATATTTATCTT 731
QY 241 ProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGluAlaLysLysValLeu 260
Db 732 CCTGAATCTCCAAGATGGCTCGTAGCAAAAGGTCGGATGCGAGAGGCAAAAGGTGTG 791
QY 261 GlnLysLeuArgGlyLysAspValSerGlyGluLeuSerLeuLeuGlyLeu 280
Db 792 CAAAAGTTACGGGGGAAAGACGATGTCAGGTGAATGTCCTTCTTCTCGAAGGGTTG 851
QY 281 GluValGlyGlyAspThrSerIleGluGluTyrIleIleGlyProAlaThrGluAlaAla 300
Db 852 GAGGTTGGAGGAGACACTTCCATTGAAGAGTACATCATTTGGACCTGCCACCGAGGAGCC 911
QY 301 AspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyrGlyProGluGly 320
Db 912 GATGATCTTGTACTGACGGTGATAGGAACAATCACTTTATGGGCTGAAGAAGGC 971
QY 321 GlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGlySerValLeuSerLeu 340
Db 972 CAGTCATGGATTGCTCGACCTTCTAAGGGACCCATCATCTTGAAGTGTGCTTCTCTT 1031
QY 341 AlaSerArgHisGlySerMetValAsnGlnSerValProLeuMetAspProIleValThr 360
Db 1032 GCATCTGTCATGGGAGCATGGTGAACAGAGTGTACCCCTTATGGATCCGATGTGACA 1091
QY 361 LeuPheGlySerValHisGluAsnMetProGlnAlaGlyGlySerMetArgSerThrLeu 380
Db 1092 CTTTTGGTAGTGTCCATGAGAATATGCCCTCAAGCTGGAGAAAGTATGAGGAGCACATTG 1151
QY 381 PheProAsnPheGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTrp 400
Db 1152 TTTCAAAACCTTGGAAAGTATGTTTCAGTGTACAGATCAGCATGCCAAAATCAGCAGTGG 1211
QY 401 AspGluGluAsnLeuHisArgAspAspGluGluTyrAlaSerAspGlyAlaGlyGlyAsp 420
Db 1212 GATCAAGAGAAATCTTCATAGGGATGACGAGGAGTACGCATCTGATGGTGAGGAGGTGAC 1271
QY 421 TyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGlyLys 440
Db 1272 TATGAGGACAATCTCCATAGCCCATTTGCTGTCCAGGCGAGGCAACAGGTGCGGAAGGAAG 1331
QY 441 AspIleValHisHisGlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeu 460
Db 1332 GACATTTGTGACCATGTCACCGTCAACGTTGAGTGTGAGCATGAGAAGGCAAAACCTCTTA 1391
QY 461 GlyGluGlyGlyAspGlyValSerSerThrAspIleGlyGlyGlyTrpGlnLeuAlaTrp 480
Db 1392 GGGGAGGTTGGAGATGTTGAGCAGCACTGATATCGGTGGGGATGGCAGCTTGTCTTGG 1451
QY 481 LysTrpSerGluLysGluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyr 500
Db 1452 AAATGGTTCAGAGAAAGGAGGTGAGAAATGTTAGAAAGGAGGTGGTTTCAAAAGAGTCTAC 1511
QY 501 LeuHisGlnGluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGly 520
Db 1512 TTGCACCAAGAGGAGTCTCTGGCTCAAGAGGGGCTCAATTGTTTCACTTCCCGGTGGT 1571
QY 521 GlyAspValPheGluGlySerGluPheValHisAlaAlaLeuValSerGlnSerAla 540
Db 1572 GCGCATGTTTTTGGAGGTAGTGTGATTTGTACATGCTGCTGCTTTTGTAGTCAAGTCAGCA 1631
QY 541 LeuPheSerLysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisProSer 560
Db 1632 CTTTTCTCAAAGGCTTGTGCTGAACCCAGCATGTGATGCTGCCATGGTTTCAACCATCT 1691

Qy 561 GluValAlaAlaLysGlySerArgTrpLysAspLeuPheGluProGlyValArgAla 580
Dd 1692 GAGGTAGCTGCCAAGGTTACGTTGGAAAGATTTGTTGAACCTGGAGTGAGCGGTGCC 1751
Qy 581 LeuLeuValGlyValGlyIleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyValLeu 600
Dd 1752 CTGTTAGTCGGTGTGGAATTCAGATCCTTCAACAGTTTCTGGAATAAACGGGTCTCTG 1811
Qy 601 TyrTyrThrProGlnIleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGly 620
Dd 1812 TACTATACCCACAAATCTTGAGCAAGCTGGTGTGGCAGTTATCTTTCCAAATTGGT 1871
Qy 621 LeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuPro 640
Dd 1872 CTCAGCTCGCATCAGCATCCATCTTGATCAGTTCTCTCACTACCTTACTAATGCTTCCT 1931
Qy 641 CysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuGly 660
Dd 1932 TGCATTGGCTTTGCCATGCTGCTTATGGATCTTCCGGAAGAGGTTTTTGTGCTAGGC 1991
Qy 661 ThrIleProIleLeuIleAlaSerLeuValIleLeuValValSerAsnLeuIleAspLeu 680
Dd 1992 ACAATTCCATCTTGATAGCATCTCTAGTTATCTCTGTTGTGTGCAATCTAATTGATTG 2051
Qy 681 GlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyrPheCysCysPhe 700
Dd 2052 GGTACACTAGCCCATGCTTTGTCTCTCCACCGTCAGTGTATCTGCTACTTCTGCTGCTC 2111
Qy 701 ValMetGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePheProThrArgVal 720
Dd 2112 GTTATGGGATTGGTCCCATCCCAACATTTATGTGCAGAGATCTTTCCAAACGAGGTT 2171
Qy 721 ArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIleGlyAspIleIleValThr 740
Dd 2172 CGTGGCCTCTGTATTGCCATTTGTGCTTACATTTCTGGATCGAGATATCATCGTCACC 2231
Qy 741 TyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAla 760
Dd 2232 TACAGCCTTCTCTGATGCTGAATGCTATTGGACTGGCGGGTGTTTTCAGCATATATGCA 2291
Qy 761 ValValCysLeuIleSerPheValPheValPheLeuLysValProGluThrLysGlyMet 780
Dd 2292 GTCGTATGCTTGATTTCCCTTTGTGTTCTGCTTCCCTTAAGGTCCTCGACAAAGGGGATG 2351
Qy 781 ProLeuGluValIleThrGluPhePheAlaValGlyAlaLysGlnAlaAlaLysAla 800
Dd 2352 CCCCTTGAGGTTATTACCGAATTTCTTTGCAGTTGGTGCGAAGCAAGCGGCTGCAAAAGCC 2411
RESULT 2
ADM47932
ID ADM47932 standard; DNA; 2908 BP.
XX
AC ADM47932;
XX
DT 03-JUN-2004 (first entry)
XX
DE Polynucleotide sequence #350 useful in producing transgenic plants.
XX
DE Plant; maize; rice; sorghum; thale cress; soybean; wheat; transgenic;
KW osmotic stress; sugar transport; cell cycle pathway; plant height;
KW carbohydrate transport; crop productivity; plant growth;
KW stress resistance; disease resistance; insect resistance; heat tolerance;
KW nitrogen assimilation; water stress tolerance;
KW photosynthetic carbon fixation; virus resistance; gene therapy; gene; ds.
XX
OS Zea mays.
XX
PN US2003233670-A1.
XX
PD 18-DEC-2003.
XX
PF 04-DEC-2002; 2002US-00310154.

PR 04-DEC-2001; 2001US-0337358P.
XX (EDGE/) EDGERTON M D.
PA (CHOM/) CHOMET P S.
PA (LACC/) LACCETTI L B.
XX
PI Edgerton MD, Chomet PS, Laccetti LB;
XX
DR WPI; 2004-061374/06.
DR P-PSDB; ADM48300.
XX
PT New polynucleotide, useful for manipulating plant protein quality,
PT improving plant growth, yield and crop productivity or grain composition
PT or producing plants with improved properties.
XX
PS Claim 1; SEQ ID NO 350; 144pp; English.
XX
CC The present invention relates to polynucleotide sequences, and the
CC proteins they encode. The sequences are isolated from a variety of
CC organisms such as plants (e.g. maize, rice, sorghum, thale cress,
CC soybean, and wheat), cyanobacteria, bacteria, yeast and other fungi. The
CC polynucleotide and polypeptide sequences of the invention are useful in
CC the production of transgenic plants that have improved properties. Also
CC disclosed are methods of producing fertile transgenic plants, preferably
CC maize, with desired phenotypes. The polynucleotide and polypeptide
CC sequences are useful for improving plants by providing protection against
CC osmotic stress, improving altering sugar transport and/or metabolism,
CC modifying the cell cycle pathway, reducing plant height, modifying
CC carbohydrate transport, improving crop productivity, improving plant
CC growth and stress resistance, improving disease resistance, improving
CC insect resistance, improving cold or heat tolerance, improving nitrogen
CC assimilation, improving stalk strength, improving water stress tolerance,
CC improving photosynthetic carbon fixation, improving biotic and abiotic
CC stress resistance, improving senescence, and conferring virus resistance.
CC The present sequence represents a polynucleotide sequence of the
CC invention. Note: The sequence data for this patent is not provided in the
CC printed specification but is obtained in electronic format from the USPTO
CC website at seqdata.uspto.gov.
XX
SQ Sequence 2908 BP; 669 A; 640 C; 747 G; 852 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 0 Length: 2908
Score: 4059.00 Matches: 797
Percent Similarity: 99.75% Conservative: 1
Best Local Similarity: 99.62% Mismatches: 2
Query Match: 99.61% Indels: 0
DB: 12 Gaps: 0
US-10-051-909-32 (1-800) x ADM47932 (1-2908)
Qy 1 IleArgSerGlySerTrpLeuAlaValGlnThrProPheThrProAspLeuAspArg 20
Dd 166 ATTCGGAGCGGCTCTTGGCTTGCAATCCAGACGCCCTTCACCCCTGATCTGGACCGGAGG 225
Qy 21 GluArgLeuLeuProSerValValLeuAlaLeuProGlyProLeuProAlaSerCys 40
Dd 226 GAGCGGCTCCTTCCGTCAGTTGTTCTTGTGTTGGCGCTCTTCCGCTGCTTGGTGT 285
Qy 41 SerSerGlnGluProValThrSerAspAspIleLeuGluAspLysMetSerGlyAlaVal 60
Dd 286 TCTTCACAGGAGCGCGGTGACCTCGGACGATATCTTGGAGGACAAGATGTTCGGGGGCTGTT 345
Qy 61 LeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrpAspAsnAlaThrIle 80
Dd 346 CTTGTGCGCCATAGTCGCCCTCCATCGGCAATCTATTGAGGGGTGGACAATGCCACCATC 405
Qy 81 AlaAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsnGluProThrValGlu 100
Dd 406 GCAGCTGCTGTTCTGTATATAAAGAAGGAATTTCAATTGCAAAATGAGCCACTGTGGAG 465
Qy 101 GlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThrThrPheSerGlyPro 120

Db 466 GGACTAATTGTCAATGTCACTTATCGCGCCACCACATCGTTACTACATTCTCCGGGCCA 525
QY 121 LeuSerAspSerIleGlyArgProMetLeuIleLeuSerSerIleLeuTyrPhePhe 140
Db 526 TTATCAGACTCGAATGGCCGACGCCCTATGCTTATCTCTCAATTCTGTACTTCTTC 585
QY 141 SerGlyLeuIleMetLeuTyrPheProAsnValTyrValLeuLeuAlaArgPheVal 160
Db 586 AGCGGCCATCATGCTATGCTCTCAATGTCTATGCTCTGCTGTGGCAGCGTTGCTA 645
QY 161 AspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyrIleSerGluIleAla 180
Db 646 GATGGATTGGTATTGGCTTGGCTGTACAGCTTGTGCCCTTTGTACATTTTCAGAAATAGCC 705
QY 181 ProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSerGlySerGlyGlyMet 200
Db 706 CCTTCGGAGATTAGAGTTTGTGTAATACACTACCAAAATTCAAGTGGATCAGGAGGAATG 765
QY 201 PheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSerProAspTyrArgIle 220
Db 766 TTCTTGTCTACTGTCATGGTGTGGATGTCCTGTGCGCATCACCCGATTGGAGAATT 825
QY 221 MetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeuThrIlePheTyrLeu 240
Db 826 ATGCTTGGTGTGCTCGCATACCTTCTTCTTCTTGGTTGACAAATATTTATCTT 885
QY 241 ProGluSerProArgTyrLeuValSerLysGlyArgMetAlaGluAlaLysLysValLeu 260
Db 886 CCTGAATCTCCAAGATGGCTCGTTAGCAAGGTCCGATGGCAGAGGCCAAAGAGTGTG 945
QY 261 GlnLysLeuArgGlyLysAspValSerGlyGluLeuSerLeuLeuLeuGluGlyLeu 280
Db 946 CAAAAGTTACGGGGGAAAGACCATGTCTCAGGTGAATTGTCCTTCTTCTCGAAGGTTG 1005
QY 281 GluValGlyGlyAspThrSerIleGluGluTyrIleIleGlyProAlaThrGluAlaAla 300
Db 1006 GAGTTGGAGGAGACACTTCCATTGAAGAGTACATCATTTGGACCTGCCACCGAGCGAGCC 1065
QY 301 AspAspLeuValThrAspGlyAspLysGluGluIleThrLeuTyrGlyProGluGluGly 320
Db 1066 GATGATCTTGTACTGACGGTGATGAAGGAACAAATCACACTTTATGGCCCTGAAGAAGGC 1125
QY 321 GlnSerTyrIleAlaArgProSerLysGlyProIleMetLeuGlySerValLeuSerLeu 340
Db 1126 CAGTCATGGATTGCTCGACCTTCCAAAGGACCCAGCATGCTTGAAGTGTGCTTCTCTT 1185
QY 341 AlaSerArgHisGlySerMetValAsnGlnSerValProLeuMetAspProIleValThr 360
Db 1186 GCATCTCGTCATGGAGCATGGTGAACCCAGAGTGTACCCCTTATGGATCCGATGTGACA 1245
QY 361 LeuPheGlySerValHisGluAsnMetProGlnAlaGlyGlySerMetArgSerThrLeu 380
Db 1246 CTTTGTGGTAGTGTCCATGAGAAATATGCCTCAAGCTGAGGAAGTATGAGGAGCACATTG 1305
QY 381 PheProAsnPheGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTyr 400
Db 1306 TTTCCAAAACCTTGGAAAGTATGTTCAAGTGTACAGATCAGCATGCCAAAATGAGCAGTGG 1365
QY 401 AspGluGluAsnLeuHisArgAspAspGluGluTyrAlaSerAspGlyAlaGlyGlyAsp 420
Db 1366 GATGAAGAGAAATCTTCATAGGATGACGAGGAGTACGCATCTGATGCTGCGAGGAGTGAC 1425
QY 421 TyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLys 440
Db 1426 TATCAGGACAAATCTCCATAGCCCATTTGCTGTCCAGGAGGCAACAGGTGCGAAGGGAAG 1485
QY 441 AspIleValHisHisGlyHisArgGlySerAlaLeuSerMetMetArgArgGlnThrLeuLeu 460
Db 1486 GACATTGTGCACCATGTTACCGTGAAGTGTGAGCATGAGAGGCAAGCCCTCTTA 1545
QY 461 GlyGluGlyGlyAspGlyValSerSerThrAspIleGlyGlyTyrPheGlnLeuAlaTyr 480

Db 1546 GGGGAGGTTGAGATGGTGTGAGCAGCACTGATATCGGTGGGGATGGCAGCTTGTCTTGG 1605
QY 481 LysTyrSerGluLysGluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyr 500
Db 1606 AAATGGTCAGAGAAGGAAGGTGAGATGGTAGAAAGGAAGTGGTTTCAAAGAGTCTAC 1665
QY 501 LeuHisGlnGluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGly 520
Db 1666 TTGCACCAAGAGGGAGTTCTCTGGCTCAAGAAGGGGCTCAATTGTTTCACTTCCCGGTGGT 1725
QY 521 GlyAspValPheGluGlySerGluPheValHisAlaAlaLeuValSerGlnSerAla 540
Db 1726 GGCGATGTTCTTGAGGGTAGTGAGTTGTACATGCTGCTGCTTAGTAAGTCAGTCAGCA 1785
QY 541 LeuPheSerLysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisProSer 560
Db 1786 CTTTCTCAAAAGGGTCTTGTCTGAACCAACGATGTTCAGATGCTGCCATGTTTCAACCATCT 1845
QY 561 GluValAlaAlaLysGlySerArgTyrLysAspLeuPheGluProGlyValArgAla 580
Db 1846 GAGGTAGTGCCTCAAGGTTTCAAGTTGGAAGATTTGTTGAACCTGGAGTGGCGGTGCC 1905
QY 581 LeuLeuValGlyValGlyIleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyValLeu 600
Db 1906 CTGTTAGTCGGTGTGGAAATTCAGATCCTTCAACAGATTTGCTGGAATAAACGGTGTCTG 1965
QY 601 TyrTyrThrProGlnIleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGly 620
Db 1966 TACTATACCCCAAAATTCCTGAGCAAGCTGGTGTGGCAGTTATTTTCCAAATTTGGT 2025
QY 621 LeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuPro 640
Db 2026 CTCAGCTCGGCATCAGCATCCATCTTGATCAGTTCTCTCACTACCTTACTAATGCTTCT 2085
QY 641 CysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuGly 660
Db 2086 TGCAATTGGCTTTGCCATGCTGCTTATGGATCTTTCCGGAAGAAGTTTGTGCTGTAGGC 2145
QY 661 ThrIleProIleLeuIleAlaSerLeuValIleLeuValValSerAsnLeuIleAspLeu 680
Db 2146 ACAATTCCAATCTTGATAGCATCTCTAGTTATCTGTTGTGTCCAATCTAATGATTG 2205
QY 681 GlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyrPheCysCysPhe 700
Db 2206 GGTACACTAGCCCATGCTTGTCTCTCCACCGTCAGTGTATCGTCTACTTCTGCTGCTTC 2265
QY 701 ValMetGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePheProThrArgVal 720
Db 2266 GTTATGGGATTTGGTCCCATCCCCAACATTTTATGTGCAGAGATCTTTCCAACAGGGTT 2325
QY 721 ArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIleGlyAspIleIleValThr 740
Db 2326 CGTGGCCTCTGTATTGCCATTGTGCCCTTTACATCTGGATCGGAGATATCATCGTCACC 2385
QY 741 TyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAla 760
Db 2386 TACAGCCTTCTCTGTGATCTGAAATGCTATTGGACTGGCGGGTGTTTTTCAGCATATATGCA 2445
QY 761 ValValCysLeuIleSerPheValPheValPheLeuLysValProGluThrLysGlyMet 780
Db 2446 GTCGTATGCTTGTATTCTTGTGTTCTCTTAAAGGTCCCTGAGACAAAGGGGATG 2505
QY 781 ProLeuGluValIleThrGluPhePheAlaValGlyAlaLysGlnAlaAlaLysAla 800
Db 2506 CCCCTTGAGGTTATTACCGAATTTCTTTGCAAGTGGTGGCAAGCAAGCGGCTGCAAAAGCC 2565
RESULT 3
ABK51962
ID ABK51962 standard; cDNA; 2824 BP.
XX
AC ABK51962;
XX
DT 27-AUG-2002 (first entry)

```
XX DE      Corn contig encoding Arabidopsis thaliana-like sugar transport protein.
XX DE
XX KW      Corn; Arabidopsis thaliana-like sugar transport protein;
XX KW      carbohydrate transport; grain filling; annual field crop; plant; gene;
XX KW      ss.
XX XX
OS Zea mays.
XX XX
FH Key      Location/Qualifiers
FT CDS      238..2481
FT          /*tag= a
FT          /product= "Corn Arabidopsis thaliana-like sugar transport
FT          protein"
XX XX
XX PN      US6383776-B1.
XX PD      07-MAY-2002.
XX XX
XX PF      14-APR-1999; 99US-00291922.
XX PR      24-APR-1998; 98US-0083044P.
XX XX
XX PA      (DUPO ) DU PONT DE NEMOURS & CO E I.
XX XX
XX PI      Allen SM, Hitz WD, Kinney AJ, Tingey SV;
XX XX
XX DR      WPI; 2002-453364/48.
XX DR      P-PSDB; AAU97201.
XX XX
XX PT      New nucleic acid encoding plant sugar-transport proteins, useful for
XX PT      preparing transgenic plants with altered carbohydrate distribution.
XX XX
XX PS      Claim 3; Col 25-28; 54pp; English.
XX CC
XX CC      The present invention relates to the isolation of plant polynucleotide
XX CC      sequences encoding an Arabidopsis thaliana-like sugar transport protein
XX CC      or Beta vulgaris-like sugar transport protein. The polynucleotide
XX CC      sequences are useful for altering the level of sugar transport proteins
XX CC      in plants, i.e. for control of carbohydrate transport and distribution in
XX CC      plant cells, e.g. during grain filling of annual field crops (e.g. corn,
XX CC      rice, soybeans, and wheat), and, for studying carbohydrate flows and
XX CC      sugar transport. The polynucleotide sequences can also be used to isolate
XX CC      cDNA sequences and genes that encode homologues of the new proteins. The
XX CC      present sequence representing a contig assembled from various corn cDNA
XX CC      clones encodes an Arabidopsis thaliana-like sugar transport protein
XX XX
XX SQ      Sequence 2824 BP; 644 A; 649 C; 745 G; 775 T; 0 U; 11 Other;

Alignment Scores:
Pred. No.:      4.19e-298      length:      2824
Score:          3517.00      Matches:      684
Percent Similarity: 95.18%      Conservative: 27
Best Local Similarity: 91.57%      Mismatches: 34
Query Match:      86.31%      Indels:      2
DB:              6      Gaps:        2

US-10-051-909-32 (1-800) x ABK51962 (1-2824)
QY 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuGlnGlyTrp 75
D 238 ATGGGGGCGCGTGATGGTCGCCATCGCGCCCTCTATCGGCAACTTGCTGCAGGGCTGG 297
QY 76 AspAsnAlaThrIleAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsn 95
D 298 GACAATGCAATGTGAGCCGCTCTGTACATAAAGAAGGAATTCAACCTGCAGAGC 357
QY 96 GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115
D 358 GAGCCTCTGATCGAAGGCCTCATCGTCGCTATGCTTCCTCATTGGGGCAACAGTCATCACA 417
QY 116 ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer 135
```

```
Db 418 ACATCTCGGGGCAAGGGCTGACTGCGTTGGTAGGAGGCCCATGCTGGTCGCCTCGGCT 477
QY 136 IleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyrValLeuLeu 155
D 478 GTCCTCTACTTCGTCAGTGGGCTGGTGAATGCTTTGGGCGCAATTGIGTACATCTTGCTC 537
QY 156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr 175
D 538 CTCGCAAGGCTCATTTGATGGGTTTCGGTATCGGTTTGGCGGTACACACTTGTCTCTCTAC 597
QY 176 IleSerGluIleAlaProSerGluIleArg--GlyLeuLeuAsnThrLeuProGlnPhe 194
D 598 ATCTCCGAAACTGCACCGCACAGANATTCTTGGGGCTGNTNGAACACGTTCCCGCAGTTC 657
QY 195 SerGly---SerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSer 213
D 658 ATTGGGGTCAGNGAGGGATGTTCTCTCTCTACTGTCATGGTGTGGGATGTCCCTCATG 717
QY 214 ProSerProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePhe 233
D 718 CCCAAACCTGATTGGAGGCTCATGCTTGGAGTTCTGTCGATCCCGTCACTTATNTACTTT 777
QY 234 GlyLeuThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMet 253
D 778 GGAAGTACTGCTCTACTTGCCTGAATCAACAAGGTGGCTTGTNAGCAAAGGAAGGATG 837
QY 254 AlaGluAlaLysLysValLeuGlnLysLeuArgGlyLysAspValSerGlyGluLeu 273
D 838 GCGGAGCGCAAGAGAGTGTGCAAGGCTCGGGGAAGAGAGATGTCTCANGGAGANG 897
QY 274 SerLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGlyTyrIleIle 293
D 898 GCTCTTCTAGTTGAAGGTTTGGGGGTCGGTAAAGATACACGTTATTNAGAGTACATCATT 957
QY 294 GlyProAlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThr 313
D 958 GGACCTGCCACCGAGGCGCGATGATCTTGAATGACGGTGATAAGGAACAATCACA 1017
QY 314 LeuTyrGlyProGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMet 333
D 1018 CTTTATGGGCTGAAGAAGGCCAGTCATGGATTGCTCGACCTTCTAAGGACCCATCATG 1077
QY 334 LeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValPro 353
D 1078 CTTTGAAGTGTGCTTCTCTTGCATCTCGTCATGGGAGCATGGTGAACACGAGTGTACCC 1137
QY 354 LeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly 373
D 1138 CTTATGGATCCGATGTGTACACTTTTGGTAGTGTCCATGAGAATATGCTCAAGCTGGA 1197
QY 374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393
D 1198 GGAAGTATGAGGAGCACATTTTCCAACTTTGGAAGTATGTTTCAGTGTACAGATCAG 1257
QY 394 HisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAla 413
D 1258 CATGCCAAAAATGACGAGTGGGATGAAGAGAATCTTCATAGGGATGACGAGGATACGCA 1317
QY 414 SerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGln 433
D 1318 TCTGATGGTGCAGGAGGTGACTATGAGGACAATCTCCATAGCCCATTTGCTGTCCAGGCAG 1377
QY 434 AlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeuSer 453
D 1378 GCAACAGGTGCGGAAGGAAGGACATTGTGCACCATGGTCAACCGTGAAGTGCTTTTGAGC 1437
QY 454 MetArgArgGlnThrLeuLeuGlyGlyGlyAspGlyValSerSerThrAspIleGly 473
D 1438 ATGAGAAGGCAAGCCTCTTAGGGGAGGGTGGAGATGGTGTGAGCAGCACTGATATCGGT 1497
QY 474 GlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArgLysGlu 493
D 1498 GGGGGATGGCAGCTTCTTGGAAATGGTCAAGAAAGGAAGGTGAGATGGTAGAAGGAA 1557
```


Qy 494 GlyGlyPheLysArgValTyrLeuHisGlnGluValProGlySerArgGlySer 513
Db 1558 GGTTGGTTTCAAAAGAGTCTACTTGCACCAAGAGGAGTCTCTGGCTCAAGAAGGGGCTCA 1617
Qy 514 IleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaAla 533
Db 1618 ATTGTTTCACTTCCCGGTGGTGGCGATGTTCTTGAAGGTAGTGAGTTGTGTACATGCTGCT 1677
Qy 534 AlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAsp 553
Db 1678 GCTTAGTAAGTCAGTCAGCACTTCTCAAAAGGTTCTGTGAACCAAGCATGTCAGAT 1737
Qy 554 AlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuPhe 573
Db 1738 GCTGCATGTTTCAACCATCTCAGGTAGCTGCCAAAGGTTACGTTGGAAGATTTGTTT 1797
Qy 574 GluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnPhe 593
Db 1798 GAACCTGGAGTGAGCGGTGCTGTAGTCGGTGTGGAATTCAGATCCTTCAACAGTTT 1857
Qy 594 AlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyValAla 613
Db 1858 GCTGGAATAAACGGTGTCTCTACTATACCCCAAAATCTTGAGCAAGCTGGTGTGGCA 1917
Qy 614 ValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeu 633
Db 1918 GTATATCTTTCCAAATTTGGTCTCAGCTCGGCATCAGCATCCATCTTGATCAGTTCTCTC 1977
Qy 634 ThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGly 653
Db 1978 ACTACCTTACTAATGCTTCCCTTGCATTTGGCTTGGCATGCTGTATGGATCTTTCCGGA 2037
Qy 654 ArgArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuVal 673
Db 2038 AGAAGGTTTTTGTGCTAGGCACAATTTCCAATCTTGATAGCATCTCTAGTATCCTGGTT 2097
Qy 674 ValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuSerThrValSerVal 693
Db 2098 GTGTCCAATCTAATTGATTGGGTACACTAGCCCATGCTTTGCTCTCCACCATCAGTGT 2157
Qy 694 IleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAla 713
Db 2158 ATCGTCTACTTCTGTGCTGCTTCTGTTATGGGATTTGGTCCCATCCCCAACATTTATGTGCA 2217
Qy 714 GluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrp 733
Db 2218 GAGATCTTTTCCAACCAAGGTTGCTGGCCTCTGTATGGCATTTGTGCCITTTACATTTCTGG 2277
Qy 734 IleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAla 753
Db 2278 ATCGGAGATATCATCGTCACTACCTACAGCCTTCTGTGATGCTGAATGCTATTGGACTGGCG 2337
Qy 754 GlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeuLys 773
Db 2338 GGTGTTTTTCAGCATATATGAGTCGTATGCTTGATTTCTTGTGTTCTGCTCTCTTAAG 2397
Qy 774 ValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAla 793
Db 2398 GTCCCTGAGACAAAGGGGATGCCCTTGAGGTTATTACCGAATTTCTTTGCAAGTTGGTGGC 2457
Qy 794 LysGlnAlaAlaAlaLysAla 800
Db 2458 AAGCAAGCGGCTGCAAAAGCC 2478

RESULT 4
ABX93198
ID ABX93198 standard; cDNA; 2824 BP.
XX
AC ABX93198;
XX
DT 29-MAY-2003 (first entry)
XX

DE CDNA encoding corn sugar transport protein #1.
XX Arabidopsis thaliana-like sugar transport protein; corn; rice; wheat;
KW plant sugar transport protein; carbohydrate transport; soybean;
KW carbohydrate distribution; plant; gene; ss.
XX Zea mays.
OS
PN US2002178468-A1.
XX
PD 28-NOV-2002.
XX
PF 17-JAN-2002; 2002US-00051902.
XX
PR 24-APR-1998; 98US-0083044P.
PR 14-APR-1999; 99US-00291922.
XX (ALLE/) ALLEN S M.
PA (HITZ/) HITZ W D.
PA (KINN/) KINNEY A J.
PA (TING/) TINGEY S V.
XX
PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;
XX WPI; 2003-340957/32.
DR P-PSDB; ABU08326.
XX
PT Novel plant sugar transport proteins and nucleic acid encoding the
PT protein useful for producing transgenic plants having altered levels of
XX sugar transport protein.
PS Claim 2; Page 13-15; 56pp; English.
XX
CC The present invention relates to the isolation of Arabidopsis thaliana-
CC like or Beta vulgaris-like sugar transport proteins, and the
CC polynucleotide sequences encoding them. The plant sugar transport
CC proteins of the invention have been isolated from corn, rice, soybean,
CC and wheat. The polypeptides of the invention may be used for altering the
CC level of expression of a sugar transport protein in a host cell, by
CC transforming a host cell with a chimeric construct encoding all, or a
CC portion of the sugar transport protein, in sense or antisense
CC orientation. Particularly, the polypeptides may provide a means to
CC control carbohydrate transport and distribution in plants. ABX93198-
CC ABX93205 represent cDNA sequences encoding Arabidopsis thaliana-like
CC sugar transport proteins
XX
SQ Sequence 2824 BP; 644 A; 649 C; 745 G; 775 T; 0 U; 11 Other;
Alignment Scores:
Pred. No.: 4.19e-298 Length: 2824
Score: 3517.00 Matches: 684
Percent Similarity: 95.18% Conservative: 27
Best Local Similarity: 91.57% Mismatches: 34
Query Match: 86.31% Indels: 2
DB: 8 Gaps: 2
US-10-051-909-32 (1-800) x ABX93198 (1-2824)
Qy 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrp 75
Db 238 ATGGGGGGCGCGGTGATGTCGCCATCGCGGCTCTATCGCAACTTGTCTGAGGGCTGG 297
Qy 76 AspAsnAlaThrIleAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsn 95
Db 298 GACAATGCGACAATTGTCGAGCGCTCTGTACATAAAGAAGGAATTCACCTGCAGAGC 357
Qy 96 GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115
Db 358 GAGCCTCTGATCGAAGGCCCTCATCGTCGCCATGTCTCATTTGGGGCAACAGTCATCACA 417
Qy 116 ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer 135
Db 418 ACATCTCCGGGGCCAAAGGGCTGACTGCGTTGGTAGAGGCCCATGCTGGTCCGCTCGGCT 477

Qy	136	IleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyrValLeuLeu	155
Db	478	GTCCCTCTACTTCGTTCAGTGGCTGGTGATGCTTTGGCGCCAATTGTGTACATTTGCTC	537
Qy	156	LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr	175
Db	538	CTCGCAAGGCTCATTTGATGGGTTTCGGTATCGGTTTGGCGGTCACACTTGTTCCTCTCTAC	597
Qy	176	IleSerGluIleAlaProSerGluIleArg--GlyLeuLeuAsnThrLeuProGlnPhe	194
Db	598	ATCTCCGAAACTGCACCGCACAGANATTCTTGGGGCTGNTNGAACACCGTTCCCGCAGTTC	657
Qy	195	SerGly--SerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSer	213
Db	658	ATTGGGGTCAGNGAGGGATGTTCTCTCCTACTGCATGGTGTGTTGGGATGTCCTCATG	717
Qy	214	ProSerProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePhePhe	233
Db	718	CCCAAACTGANTGGAGGCTCATGCTTGGAGTTCTGTGATCCCGTCACCTTATNTACTTT	777
Qy	234	GlyLeuThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMet	253
Db	778	GGACTGACTGTCTTACTTGCCTGAATCACCAAGGTGGCTTGTNAGCAAAAGGAGGATG	837
Qy	254	AlaGluAlaLysLysValLeuGlnLysLeuArgGlyLysAspValSerGlyGluLeu	273
Db	838	GCGGAGGCGAAGAGAGTGNTGCAAGGCTGCGGGAAAGAGAAGATGTTCTCANGGAGANG	897
Qy	274	SerLeuLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIle	293
Db	898	GCTCTTCTAGTTGAAGGTTTGGGGTTCGGTAAAGATACACGTATTTNAGAGTACATCATT	957
Qy	294	GlyProAlaThrGluAlaAlaAspLeuValThrAspGlyAspLysGluGlnIleThr	313
Db	958	GGACCTGCCACCGAGGCAGCCGATGATCTTGTAACTGACGGTGATAAGGAACAATAACACA	1017
Qy	314	LeuTyrGlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMet	333
Db	1018	CTTTATGGGCCTGAAGAAGGCCAGTCATGGATTGCTCGACCTTCTAAGGGACCCATCATG	1077
Qy	334	LeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValPro	353
Db	1078	CTTGGAAAGTGCTTTCTCTTGCACTCTCGTCATGGGAGCATGGTGAACACAGAGTGTACCC	1137
Qy	354	LeuMetAspProfileValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly	373
Db	1138	CTTATGGATCCGATTGTGACACTTTTGGTAGTGTCATGAGAATATGCCTCAAGCTGGA	1197
Qy	374	GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln	393
Db	1198	GGAAGTATGAGAGACACATTGTTTCCAAACTTTGGAAGTATGTTTCAGTGTCAAGATCAG	1257
Qy	394	HisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAla	413
Db	1258	CATGCCAAAAATGACGAGTGGGATGAAGAGAAATCTTCATAGGGATGACGAGGAGTACGCA	1317
Qy	414	SerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGln	433
Db	1318	TCTGATGGTGCAGGAGGTGACTATGAGGACAACTCCCATAGCCCCATTGCTGTCCAGGCAG	1377
Qy	434	AlaThrGlyAlaGluGlyLysAspIleValHisGlyHisArgGlySerAlaLeuSer	453
Db	1378	GCAACAGGTGCGGAAGGGAAGGACATTGTGCAACCATGGTCAACCGTGAAGTGTGTTGAGC	1437
Qy	454	MetArgArgGlnThrLeuLeuGlyGlyGlyAspGlyValSerSerThrAspIleGly	473
Db	1438	ATGAGAAGGCAAGCCCTCTTAGGGGAGGTGGAGATGGTGTGAGCAGCACATATCGGT	1497
Qy	474	GlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArgLysGlu	493
Db	1498	GGGGGATGGCAGCTTGCTTGGAAATGGTCAGAGAAGGAAGGTGAGAAATGGTAGAAAGGAA	1557

Qy	494	GlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgGlySer	513
Db	1558	GGTGGTTTCAAAGAGTCTACTTGCACCAAGAGGGAGTTCTCTGGCTCAAGAAGGGGCTCA	1617
Qy	514	IleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaAla	533
Db	1618	ATTGTTTTCACTTCCCGGTGGTGGCGATGTTCTTGAGGGTAGTGAGTTTGTACATGCTGCT	1677
Qy	534	AlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAsp	553
Db	1678	GCTTTAGTAAGTCAGTCAGCACTTTCTCAAAGGGTCTTGCTGAACACCGCATGTCCAGAT	1737
Qy	554	AlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuPhe	573
Db	1738	GCTGCCATGGTTTCAACCCATCTGAGGTAGTGCACAAAGGTTCCAGTTTGAAAGATTGTGTTT	1797
Qy	574	GluProGlyValArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnPhe	593
Db	1798	GAACCTGGAGTGAGGCGTGCCCTGTAGTCGGTGTGGAAATTCAGATCCTTCAACAGTTT	1857
Qy	594	AlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyValAla	613
Db	1858	GCTGGAATAAACGGTGTCTCTGTAATAACCCACAAATCTTGAGCAAGCTGGTGTGGCA	1917
Qy	614	ValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeu	633
Db	1918	GTTATTCTTTCCAAATTTGGTCTCAGCTCGGCATCAGCATCCATCCTTGATCAGTTCTCTC	1977
Qy	634	ThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGly	653
Db	1978	ACTACCTTACTAATGCTTCCCTTGCAATGGCTTTGGCATGCTGCTTATGGATCTTCCGGA	2037
Qy	654	ArgArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuVal	673
Db	2038	AGAAGGTTTTTGTCTAGGCACAAATTCCAATCTTGATAGCATCTCTAGTTATCTCTGGTT	2097
Qy	674	ValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerVal	693
Db	2098	GTGTCCAATCTAATTGATTTGGGTACACTAGCCCATGCTTTGCTCTCCACCATCAGTGTT	2157
Qy	694	IleValTyrPheCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAla	713
Db	2158	ATCGTCTACTTCTGCTGCTTCGTTATGGGATTTGGTCCCATCCCAACATTTTATGTGCA	2217
Qy	714	GluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrp	733
Db	2218	GAGATCTTTCCAAACCAAGGGTTCGTGGCCTCTGTATTGCCATTTGTGCCTTTACATTTCTGG	2277
Qy	734	IleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAla	753
Db	2278	ATCGGAGATATCATCGTCACCTACAGCCTTCTCTGTGATGCTGATGCTATTTGGACTGGCG	2337
Qy	754	GlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeuLys	773
Db	2338	GGTGTTTTTCAGCATATATGCAGTCGTATGCTTGATTTTCCCTTTGTGTTCTCTCCCTTAAG	2397
Qy	774	ValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAla	793
Db	2398	GTCCCTGAGACAAAGGGGATGCCCTTGAGGTTATTACCGAATTTCTTTGCAGTTTGGTGCG	2457
Qy	794	LysGlnAlaAlaAlaLysAla 800	
Db	2458	AAGCAAGCGGCTGCAAAAGCC 2478	

RESULT 5

ADG47905

ID ADG47905 standard; cDNA; 2824 BP.

XX

AC ADG47905;

DT 11-MAR-2004 (first entry)

XX Arabidopsis-like sugar transport protein;
 KW Beta-vulgaris-like sugar transport protein; transgenic; physical mapping;
 KW corn; plant; gene; ss.
 XX
 OS Zea mays.

XX Key Location/Qualifiers
 FH 238..2481
 FT /*tag= a
 FT /product= "Corn Arabidopsis-like sugar transport protein"
 FT /transl_except= (pos:622..624, aa:Xaa)
 FT /note= "Xaa corresponds to an unknown amino acid"
 FT /transl_except= (pos:634..636, aa:Xaa)
 FT /note= "Xaa corresponds to an unknown amino acid"
 FT /transl_except= (pos:637..639, aa:Xaa)
 FT /note= "Xaa corresponds to an unknown amino acid"
 FT /transl_except= (pos:667..669, aa:Xaa)
 FT /note= "Xaa corresponds to an unknown amino acid"
 FT /transl_except= (pos:769..771, aa:Xaa)
 FT /note= "Xaa corresponds to an unknown amino acid"
 FT /transl_except= (pos:820..822, aa:Val)
 FT /transl_except= (pos:856..858, aa:Xaa)
 FT /note= "Xaa corresponds to an unknown amino acid"
 FT /transl_except= (pos:889..891, aa:Xaa)
 FT /note= "Xaa corresponds to an unknown amino acid"
 FT /transl_except= (pos:895..897, aa:Xaa)
 FT /note= "Xaa corresponds to an unknown amino acid"
 FT /transl_except= (pos:943..945, aa:Xaa)
 FT /note= "Xaa corresponds to an unknown amino acid"

XX US2002199217-A1.

PN 26-DEC-2002.

XX 17-JAN-2002; 2002US-00051909.

XX 24-APR-1998; 98US-0083044P.

PR 14-APR-1999; 99US-00291922.

XX (HELE/) HELENTJARIS T G.

XX Helentjaris TG;

XX WPI; 2004-040967/04.

DR P-PSDB; ADG47906.

XX New isolated polynucleotide encoding a polypeptide having sugar transport protein activity, for producing a transformed plant and for use as probes in physical mapping.

XX Disclosure; SEQ ID NO 1; 71pp; English.

XX The invention relates to Arabidopsis-like or Beta-vulgaris-like sugar transport proteins and their corresponding nucleic acid sequences. The sequences of the invention are useful to transform a cell. These are also useful to produce a transgenic plant. Probes derived from sequences encoding sugar transport protein may be used for physical mapping. The present sequence is corn Arabidopsis-like sugar transport protein cDNA.

XX Sequence 2824 BP; 644 A; 649 C; 745 G; 775 T; 0 U; 11 Other;

XX Alignment Scores:

Pred. No.:	4.19e-298	Length:	2824
Score:	3517.00	Matches:	684
Percent Similarity:	95.18%	Conservative:	27
Best Local Similarity:	91.57%	Mismatches:	34
Query Match:	86.31%	Indels:	2
DB:	12	Gaps:	2

US-10-051-909-32 (1-800) x ADG47905 (1-2824)

Qy 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrp 75

238	ATGGGGGGCGCGTGTGTCGCCATCGCGCGCTCTATCGGCAACTTGTGAGGGCTGG	297
76	AspAsnAlaThrIleAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsn	95
298	GACAATGCGACAATTGCTGGAGCGCTCCTGTACATAAAGAGGAATTCAACCTGCAGAGC	357
96	GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr	115
358	GAGCCTCTGATCGAAGGCTCATCGTCGCATGTCTCTCATTTGGGGCAACAGTCATCACA	417
116	ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer	135
418	ACATCTCCGGGGCCAAAGGCTGACTCGTTGGTAGAGGCCCATGCTGGTCCGCTCGGCT	477
136	IleLeuTyrPhePheSerGlyLeuIleMetLeuTyrPheProAsnValTyrValLeuLeu	155
478	GTCTCTACTTCGTGCTGGGCTGGTATGCTTGGGGCCCAATTGTGTACATCTTGCTC	537
156	LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr	175
538	CTCGCAAGGCTCATGTATGGGTTCCGTATCGGTTGGCGGTACACACTTGTCTCTCTAC	597
176	IleSerGluIleAlaProSerGluIleArg---GlyLeuLeuAsnThrLeuProGlnPhe	194
598	ATCTCCGAAACTGCACCCGACAGANATTCTTGGGGCTGNTNGAACACCGTGGCGAGTTC	657
195	SerGly---SerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSer	213
658	ATTGGGGTCAGNGGAGGATGTTCTCTCTCTACTGATGGTGTGGGATGTCTCTCATG	717
214	ProSerProAspTyrArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePhePhe	233
718	CCCAAACTGATTGGAGGCTCATGCTTGGAGTTCTGTGATCCCGTCACTTATNTACTTT	777
234	GlyLeuThrIlePheTyrLeuProGluSerProArgTyrLeuValSerLysGlyArgMet	253
778	GGACTGACTGTCTTCTACTTGCCTGAATCACCAAGTGGCTTGTNAGCAAGAGGATG	837
254	AlaGluAlaLysLysValLeuGlnLysLeuArgGlyLysAspValSerGlyGluLeu	273
838	GCGGAGCGGAAGAGAGTGTGTCAAAGGCTGCGGGGAAGAGAGATGTCTCANGGAGANG	897
274	SerLeuLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluTyrIleIle	293
898	GCTCTTCTAGTTGAAGTTTGGGGTTCGGTAAAGATACACGATATTNAGATACATCATT	957
294	GlyProAlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThr	313
958	GGACCTGCCACCGAGGCGAGCGATGATCTTGTAACTGACGGTGATAAGGAACAAATCACA	1017
314	LeuTyrGlyProGluGluGlyGlnSerTyrIleAlaArgProSerLysGlyProIleMet	333
1018	CTTTATGGGCTGAAGAGGCGAGTCAATGGATTGTCTGACCTTCTAAGGGACCATCATG	1077
334	LeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValPro	353
1078	CTTGGAAAGTGTGCTTCTCTGTCATCTCGTCATGGAGCATGGTGAACCATGTATCCC	1137
354	LeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly	373
1138	CTTATGGATCCGATTGTGACACTTTTGGTAGTGTCCATGAGATATGCTCAAGCTGGA	1197
374	GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln	393
1198	GGAACTATGAGGAGCACATGTTTCCAAACTTTTGGAAAGTATGTTTCAGTGTACAGATCAG	1257
394	HisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluTyrAla	413
1258	CATGCCAAAAATGAGCAGTGGATGAAGAGAAATCTTCATAGGGATGACGAGGATACGCA	1317
414	SerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGln	433

Db 1318 TCTGATGGTGCAGGAGGTGACTATGAGGACAATCTCCATAGCCCATTTGCTGTCCAGGCAG 1377

QY 434 AlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeuSer 453

Db 1378 GCAACAGGTGCGGAAGGGAAGGACATTTGTACCATGTTGACCGTGGAGTGTCTTGTAGC 1437

QY 454 MetArgArgGlnThrLeuLeuGlyGluGlyGlyAspGlyValSerSerThrAspIleGly 473

Db 1438 ATGAGAAGGCAAGCCCTCTTAGGGAGGGTGGAGATGGTGTGAGCAGCACATGATATCGGT 1497

QY 474 GlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArgLysGlu 493

Db 1498 GGGGGATGGCAGCTTCTTGGAATGGTTCAGAGAAGGAGTGGAGATGGTAGAAGGAA 1557

QY 494 GlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArgGlySer 513

Db 1558 GGTGGTTTCAAAGAGTCTACTTCACCAAGAGGGAGTTCTTGCTCAAGAGGGGCTCA 1617

QY 514 IleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaAla 533

Db 1618 ATTGTTTCACCTCCCGGTGGTGGCATGTTCTTGAGGGTAGTGAGTTTGTACATGTGCT 1677

QY 534 AlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAsp 553

Db 1678 GCCTTAGTAAGTCAGTCAGCACCTTCTCAAAGGGTCTTGCTGAACCCAGCATGTTCAGAT 1737

QY 554 AlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuPhe 573

Db 1738 GCTGCCATGGTTTCAACCATCTGAGTAGTGCCTGAGGTTTCAAGTTTGGAAAGATTGTTT 1797

QY 574 GluProGlyValArgArgAlaLeuValGlyValGlyIleGlnIleLeuGlnGlnPhe 593

Db 1798 GAACCTGGAGTGAGGCGTGCCCTGTGTAGTCGGTGTGGAAATTCAGATCCTTCAACAGTTT 1857

QY 594 AlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyValAla 613

Db 1858 GCTGGAATAAACGGTGTCTGTACTATACCCCAAAATTTCTTGAGCAAGCTGGTGTGGCA 1917

QY 614 ValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeu 633

Db 1918 GTTATTCTTTCCAAATTTGGTCTCAGCTCGGCATCAGCATCCATCTTGATCAGTTCTCTC 1977

QY 634 ThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGly 653

Db 1978 ACTACCTTACTAATGCTTCCCTTGCAATGGCTTTGCCATGCTGCTTATGGATCTTTCCGGA 2037

QY 654 ArgArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuVal 673

Db 2038 AGAAGGTTTTTGTCTGTAGGCACAATCCAAATCTTGATAGCATCTCTAGTTATCCTGGTT 2097

QY 674 ValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerVal 693

Db 2098 GTGTCCCATCTAATTGATTTGGGTACTAGCCCATGCTTTGCTCTCCACCATCAGTGT 2157

QY 694 IleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAla 713

Db 2158 ATCGTCTACTTCTGCTGCTTCTGTTATGGGATTTGGTCCCATCCCAACATTTTATGTGCA 2217

QY 714 GluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrp 733

Db 2218 GAGATCTTTCCAAACCAGGGTTGGTGGCTCTGTATTGGCAATTTGTGCCTTTACATTTCTGG 2277

QY 734 IleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAla 753

Db 2278 ATCGGAGATATCATCTGTCACCTACAGCCTTCTGTGATGCTGAATGCTATTGGACGTGGCG 2337

QY 754 GlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeuLys 773

Db 2338 GGTGTTTTTCAGCATATATGCAATGCTGCTGATGCTTGTATTTCCCTTTGTGTTCTTCTTAAG 2397

QY 774 ValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAla 793

Db 2398 GTCCCTGAGACAAAGGGGATGCCCTTGAGGTTATTACCGAATTTCTTTGCAGTTGGTGCG 2457

QY 794 LysGlnAlaAlaAlaLysAla 800

Db 2458 AAGCAAGCGCTGCAAAAGCC 2478

RESULT 6

ABK51965

ID ABK51965 standard; cDNA; 2601 BP.

XX

AC ABK51965;

XX 27-AUG-2002 (first entry)

DT

XX Soybean contig encoding A. thaliana-like sugar transport protein.

DE

XX Soybean; Arabidopsis thaliana-like sugar transport protein;

KW carbohydrate transport; grain filling; annual field crop; plant; gene;

KW ss.

XX Glycine max.

OS

XX

FH Key

FT CDS

FT 175..2388

FT /*tag= a

FT /product= "Soybean Arabidopsis thaliana-like sugar

FT transport protein"

XX

XX US6383776-B1.

PN

XX 07-MAY-2002.

PD

XX 14-APR-1999; 99US-00291922.

PF

XX 24-APR-1998; 98US-0083044P.

PR

XX (DUPO) DU PONT DE NEMOURS & CO E I.

PA

XX Allen SM, Hitz WD, Kinney AJ, Tingey SV;

PI

XX WPI; 2002-453364/48.

DR

DR P-PSDB; AAU97204.

XX

PT New nucleic acid encoding plant sugar-transport proteins, useful for

PT preparing transgenic plants with altered carbohydrate distribution.

XX

PS Claim 3; Col 35-38; 54pp; English.

XX

CC The present invention relates to the isolation of plant polynucleotide

CC sequences encoding an Arabidopsis thaliana-like sugar transport protein

CC or Beta vulgaris-like sugar transport protein. The polynucleotide

CC sequences are useful for altering the level of sugar transport proteins

CC in plants, i.e. for control of carbohydrate transport and distribution in

CC plant cells, e.g. during grain filling of annual field crops (e.g. corn,

CC rice, soybeans, and wheat), and, for studying carbohydrate flows and

CC sugar transport. The polynucleotide sequences can also be used to isolate

CC cDNA sequences and genes that encode homologues of the new proteins. The

CC present sequence representing a contig assembled from various soybean

CC cDNA clones encodes an Arabidopsis thaliana-like sugar transport protein

XX

SQ Sequence 2601 BP; 648 A; 543 C; 644 G; 766 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 3.54e-224 Length: 2601

Score: 2674.00 Matches: 522

Percent Similarity: 82.20% Conservative: 92

Best Local Similarity: 69.88% Mismatches: 115

Query Match: 65.62% Indels: 18

DB: 6 Gaps: 8

US-10-051-909-32 (1-800) x ABK51965 (1-2601)

QY 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrp 75

Db 175 ATGAAAGGTGCGCTCTTGTGCTATTGCGCGCTTCCATTGGTAATTTCTCTCAAGGATGG 234
Qy 76 AspAsnAlaThrIleAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsn 95
Db 235 GATAATGCTACCATCGCGGGCTAATGTTACATTAAGAAAGACCTTGTCTTGGGAACA 294
Qy 96 GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115
Db 295 -----ACTATGAAAGGCTTGTGTGGCATGTCTCCCTGATTGGAGCAACGGTAATCACC 348
Qy 116 ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer 135
Db 349 ACATGCTCTGCTCTATAGCGGATTGGCTCGGTGCGGACCCCATGATGATAATCTCATCT 408
Qy 136 IleLeuTyrPhePheSerGlyLeuIleMetLeuTyrSerProAsnValTyrValLeuLeu 155
Db 409 GTGCTCTATTCTTGGGTGGTTGGTGATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 468
Qy 156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr 175
Db 469 TTGGCGAGGCTACTTGTATGGATTGGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 528
Qy 176 IleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSer 195
Db 529 ATATCTGAACCGCGCGCTCTGAATAAGGGGGTCTGTGAATACGCTTCTCAGTTTCAGT 588
Qy 196 GlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSer 215
Db 589 GGCTCTGGAGGAATGTTTGTGCTGCTACTGTATGGTTTGTGGCATGTCATGAGTCCCGCG 648
Qy 216 ProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeu 235
Db 649 CTTAGCTGGAGGCTCATGCTTGGGGTCTGCTCTATCTCTCTCTCTCTCTCTCTCTCTCT 708
Qy 236 ThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGlu 255
Db 709 ACCATTTTCTTGGCCGAGTCTCTCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 768
Qy 256 AlaLysLysValLeuGlnLysLeuArgGlyLysAspValSerGlyGluLeuSerLeu 275
Db 769 GCTAAGAAGGTGCTCCAAAGATTGCGGGAAGGAGGAGTGTGTGAGGAGATGGCATTG 828
Qy 276 LeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIleGlyPro 295
Db 829 CTGCTGAAGGTCTCGGATGGGGTGGGTGATACATCTATCGAAGAGTACATAATTTGGCCCT 888
Qy 296 AlaThrGluAlaAlaAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyr 315
Db 889 GCTGACGATGTGGCTGATGTCATGAACATGCAACAGAGAAAGATAAAATTCGATTATAT 948
Qy 316 GlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGly 335
Db 949 GGATCCCAAGCAGGCTTCTTGGTTATCAAAACCTGTCACTGGACAGAGTCTATTGGC 1008
Qy 336 SerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValProLeuMet 355
Db 1009 -----CTTGGCTCACCATGGAAGCATCATCAACCAAGCATGCCCTCATG 1056
Qy 356 AspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly----- 373
Db 1057 GATCCTCTGGTGACACTGTTGGTAGCATTCATGAGAAGCTCCCCGAGACAGGCAAGA 1116
Qy 374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393
Db 1117 GGAAGCATGCGAAGCACTCTGTTTCCAAATTTTGGAAAGCATGTTTCAGCACTGCTGAGCCG 1176
Qy 394 HisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAla 413
Db 1177 CATGCTAAATTAACCAATGGGATGAAGAAAGCTTACAAAGGGAACCTGAGGACTACATG 1236
Qy 414 SerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGln 433
Db 1237 TCAGATGCAACCCGTGGGACTCCGATGATAATTTGCACAGTCTCTTAATCTCACGCCAA 1296

Qy 434 AlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeu--- 452
Db 1297 ACAACAAGCCTTGAA---AAAGACTTACCTCCTCCTCCTTCCCATTGGCAGTATCCTTGGC 1353
Qy 453 SerMetArgArgGlnThrLeuLeuGlyGluGly---GlyAspGlyValSerSerThrAsp 471
Db 1354 AGCATGAGGCGTCAAGTAGTCTCATGCAAGGGTCAGGTGAGCAAGGTGGTAGTACAGGT 1413
Qy 472 IleGlyGlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArg 491
Db 1414 ATTGGTGGTGGCTGGCAACTGGCATGGAAATGGACTGATAAA---GGTGAGATGGAAAA 1470
Qy 492 LysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArg 511
Db 1471 CAACAAGGAGGTTTAAAGAGATTATTATCATGAGGAGGAGTTTCTGCACTCTCGTCGT 1530
Qy 512 GlySerIleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHis 531
Db 1531 GGATCCATTGTATCGATTCCCGGTGAAGGC-----GAATTGTCCAG 1572
Qy 532 AlaAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMet 551
Db 1573 GTGCTGCCTTGGTAAGCCAAACCCGCTCTTACTCAAGGAGCTTATTGATGGACACCCA 1632
Qy 552 SerAspAlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAsp 571
Db 1633 GTTGGGCTGCAATGGTTCAACCATCTGAGACAGCTTCAAAGGGGCCAAGTTGGAAGCT 1692
Qy 572 LeuPheGluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGln 591
Db 1693 CTTCTTGAACCAAGGTTAAGCATGTCATTGGTTGTTGGAGTTGGAATACAAATACTTTCAG 1752
Qy 592 GlnPheAlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluAlaGly 611
Db 1753 CAGTTTTCAGGGATAAATGGGTTCTATATTACACACCTCAATCCTTGAAGAGGCCGCT 1812
Qy 612 ValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSer 631
Db 1813 GTTGAAGTTCTTCTTTCAGATATAGGCAATGGCTCAGAGTCGGCATCATCTTATCAGT 1872
Qy 632 SerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeu 651
Db 1873 GCTTTCACAACCTTCTTGATGCTTCCCTGTATAGCGGTAGCCATGAAGCTCATGGATGT 1932
Qy 652 SerGlyArgArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIle 671
Db 1933 TCAGGCAGAAGGCGAGTTGCTACTTACTACATCCCGTCTGATTGTGTGCTCATATT 1992
Qy 672 LeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrVal 691
Db 1993 TTGGTCATTGGAAGCCCTGGTAATTTTGGCAATGTGCCCCATGCAGCAATCTCAACAGTA 2052
Qy 692 SerValIleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeu 711
Db 2053 TCGGTTGTGGTTTATTCTGCTGCTTGTGATGGGTTATGGACCAATTCCAACATCCTT 2112
Qy 712 CysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThr 731
Db 2113 TCCTCAGAGATTTTCCCACTAGGTGGTGGCTCTGCTATCTGCTGCTGCTGCTGCTGCT 2172
Qy 732 PheTrpIleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGly 751
Db 2173 TTCTGGATTGGAGACATCATCATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2232
Qy 752 LeuAlaGlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPhe 771
Db 2233 CTTGGTGGTGTATTCCGCATTTACGCAGTTTGTGTTTTCATCTCGTGGATATTGTTGTTT 2292
Qy 772 LeuLysValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaVal 791
Db 2293 TTGAAGGTTCCAGAAACAAAGGGCATGCCCCCTTGAAGTCATCTCTGAATCTTTTCTGTT 2352

QY 792 GlyAlaLysGlnAlaAlaAla 798
 Db 2353 GGAGCAAGCAGGCTGCTTCT 2373

RESULT 7

ABX93201
 ID ABX93201 standard; cDNA; 2601 BP.
 XX AC ABX93201;
 XX DT 29-MAY-2003 (first entry)
 XX DE cDNA encoding soybean sugar transport protein #1.

XX Arabidopsis thaliana-like sugar transport protein; corn; rice; wheat;
 KW plant sugar transport protein; carbohydrate transport; soybean;
 KW carbohydrate distribution; plant; gene; ss.

XX Glycine max.
 OS
 XX US2002178468-A1.
 PN
 XX 28-NOV-2002.
 PD
 XX 17-JAN-2002; 2002US-00051902.
 PF
 XX 24-APR-1998; 98US-0083044P.
 PR
 XX 14-APR-1999; 99US-00291922.
 XX

PA (ALLE/) ALLEN S M.
 PA (HITZ/) HITZ W D.
 PA (KINN/) KINNEY A J.
 PA (TING/) TINGEY S V.

XX Allen SM, Hitz WD, Kinney AJ, Tingey SV;
 PI WPI; 2003-340957/32.
 XX P-PSDB; ABU08329.
 DR

XX Novel plant sugar transport proteins and nucleic acid encoding the
 PT protein useful for producing transgenic plants having altered levels of
 PT sugar transport protein.

XX Claim 2; Page 19-20; 56pp; English.

XX The present invention relates to the isolation of Arabidopsis thaliana-
 CC like or Beta vulgaris-like sugar transport proteins, and the
 CC polynucleotide sequences encoding them. The plant sugar transport
 CC proteins of the invention have been isolated from corn, rice, soybean,
 CC and wheat. The polypeptides of the invention may be used for altering the
 CC level of expression of a sugar transport protein in a host cell, by
 CC transforming a host cell with a chimeric construct encoding all, or a
 CC portion of the sugar transport protein, in sense or antisense
 CC orientation. Particularly, the polypeptides may provide a means to
 CC control carbohydrate transport and distribution in plants. ABX93198-
 CC ABX93205 represent cDNA sequences encoding Arabidopsis thaliana-like
 CC sugar transport proteins

XX SQ Sequence 2601 BP; 648 A; 543 C; 644 G; 766 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	3.54e-224	Length:	2601
Score:	2674.00	Matches:	522
Percent Similarity:	82.20%	Conservative:	92
Best Local Similarity:	69.88%	Mismatches:	115
Query Match:	65.62%	Indels:	18
DB:	8	Gaps:	8

US-10-051-909-32 (1-800) x ABX93201 (1-2601)

QY 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrp 75
 Db 175 ATGAAAGGTGCCGTCCTGTTGCTATTGCGCTTCCATTGGTAATTCCTCCAAAGGATGG 234

QY	76	AspAsnAlaThrIleAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsn	95
Db	235	GATAATGCTACCATGCGCGGGCTAATGGTTACATTAAGAAAGACCTTGTCTTGGGAACA	294
QY	96	GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr	115
Db	295	-----ACTATGGAAGGCTTGTGGTGGCATGTCCCTGATGGAGCAACGGTAATCACC	348
QY	116	ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer	135
Db	349	ACATGCTCTGCTCTATAGCGGATTGGCTCGGTGGCGGCCCATGATGATAATCTCATCT	408
QY	136	IleLeuTyrPhePheSerGlyLeuIleMetLeuTyrSerProAsnValTyrValLeuLeu	155
Db	409	GTGCTCTATTCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	468
QY	156	LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr	175
Db	469	TTGGCGAGGCTACTTGTATGGATTGGGATTGGCTTGGCTTGGTGGCTTGTCCCGGCTAT	528
QY	176	IleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSer	195
Db	529	ATATCTGAAACGGCGCGCTCTGAAATAAGGGGGTGGTGGTGGTGGTGGTGGTGGTGGT	588
QY	196	GlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSer	215
Db	589	GGCTCTGGAGGAATGTTTTTGTCTGCTACTGTATGGTTTTTGGCATGTCTATTGATCCCGG	648
QY	216	ProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeu	235
Db	649	CCTAGCTGGAGGCTCATGCTTGGGGTCTGTCTATTCTCTCTCTCTCTCTCTCTCTCTCT	708
QY	236	ThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGlu	255
Db	709	ACCATTTTTTCTTGGCGAGTCTCTCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	768
QY	256	AlaLysLysValLeuGlnLysLeuArgGlyLysAspAspValSerGlyGluLeuSerLeu	275
Db	769	GCTAAGAGGTGCTCCAAAGATTGGCGGAAGGGAGGATGTGTGAGGAGATGSCATTG	828
QY	276	LeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIleGlyPro	295
Db	829	CTGGTTGAAGGTCTCGGGATTGGGGGTGATACATCATCGAAGAGTACATAATTGGCCCT	888
QY	296	AlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyr	315
Db	889	GCTGACGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	948
QY	316	GlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGly	335
Db	949	GGATCCCAAGCAGGCGCTTCTTGGTTATCAAAACCTGTCTACTGGACAGAGTCTATTGGC	1008
QY	336	SerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValProLeuMet	355
Db	1009	-----CTTGGCTCACCATCGAAGCATCATCAACCAAGCATGCCCCCTCATG	1056
QY	356	AspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly-----	373
Db	1057	GATCCTCTGGTGACACTGTTTGGTAGCATTCATGAGAAAGCTCCCCGAGACAGGACAAGA	1116
QY	374	GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln	393
Db	1117	GGAAGCATGCGAAGCACTCTGTTCCAAATTTTGGAAAGCATGTTTCAGCACTGCTGAGCCG	1176
QY	394	HisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAla	413
Db	1177	CATGCTAAATTTGAACAATGGGATGAAGAAAGCTTACAAAGGGAACCTGAGGACTACATG	1236
QY	414	SerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGln	433
Db	1237	TCAGATGCAACCCGTTGGGGACTCCGATGATATAATTTCACAGTCTCTTTAATCTCACGCCAA	1296

QY 96 GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115
Db 295 -----ACTATGGAAGGCTTGTTGGTGGCATGTCCTGATGGAGCAACGGTAATCACC 348
QY 116 ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer 135
Db 349 ACATGCTCTGGTCTCTATAGCGGATGGCTCGGTGGCGACCCATGATGATAATCTCATCT 408
QY 136 IleLeuTyrPhePheSerGlyLeuIleMetLeuTyrPheSerProAsnValTyrValLeuLeu 155
Db 409 GTGCTCTATTTCTTGGGTGGTTGGTGATGCTGTGGTCCCAAAATGTGATGTGTGTGC 468
QY 156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr 175
Db 469 TTGGCGAGGCTACTTGATGGATTGGGATTGGCTTGTCTGTGACTCTTGTCCCGGCTCTAT 528
QY 176 IleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSer 195
Db 529 ATATCTGAACGGCGCGCTCTGAATAAGGGGTGCTTGAATACGCTTCCTCAGTTCAGT 588
QY 196 GlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSer 215
Db 589 GGCTCTGGAGGAATGTTTTGTCTGACTGTATGGTTTTTGGCATGTCTATTGATTCCTG 648
QY 216 ProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeu 235
Db 649 CCTAGCTGGAGGCTCATGCTTGGGGTCTGTCTATTCTTCTCTCTTTGTCATTG 708
QY 236 ThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGlu 255
Db 709 ACCATTTTTTCTTGTCCCGAGTCTCCTCGGTGGCTGGTCAGCAAGGAAGGATGCTCGAG 768
QY 256 AlaLysLysValLeuGlnLysLeuArgGlyLysAspValSerGlyGluLeuSerLeu 275
Db 769 GCTAAGAAGGTGCTCCAAAGATTGCGCGAAGGAGGATGTGTGAGGAGATGATGATATAT 828
QY 276 LeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIleGlyPro 295
Db 829 CTGGTTGAAGGTCTCGGATTGGGGTGATACATCTATCGAAGAGTACATAATTGGCCCT 888
QY 296 AlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyr 315
Db 889 GCTGACGATGTGGCTGATGTCATGAACATGCAACAGAGAAAGATAAAATTCGATTATAT 948
QY 316 GlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGly 335
Db 949 GGATCCCAAGCAGGCTTCTTGGTTATCAAAAACCTGTCACTGGACAGAGTCTTATTGGC 1008
QY 336 SerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValProLeuMet 355
Db 1009 -----CTTGGTCCACCATGGAAGCATCATCAACCAAGCATGCCCTCATG 1056
QY 356 AspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly----- 373
Db 1057 GATCCTCTGGTGACACTGTTTGGTAGCATTCATGAGAAGCTCCCCCGACAGAGCAAGA 1116
QY 374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393
Db 1117 GGAAGCATGCGAAGCACTCTGTTTCCAAATTTTGAAGCATGTTTCAGCACTGCTGAGCCG 1176
QY 394 HisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAla 413
Db 1177 CATGCTAAAAATTGAACAATGGGATGAAGAAAGCTTACAAAGGGAACGTGAGGACTACATG 1236
QY 414 SerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGln 433
Db 1237 TCAGATGCAACCCGTGGGACTCCGATGATAATTGACAGTCTCTTAATCTCACGCCAA 1296
QY 434 AlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeu--- 452
Db 1297 ACAACAAGCCTTGAA---AAAGACTTACCTCTCCTCCTTCCCATGGCAGTATCCTTGGC 1353

QY 453 SerMetArgArgGlnThrLeuLeuGlyGluGly---GlyAspGlyValSerSerThrAsp 471
Db 1354 AGCATGAGCGCTCACAGTAGTCTCATGCAAGGGTCAGGTGAGCAAGGTGGTAGTACAGGT 1413
QY 472 IleGlyGlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArg 491
Db 1414 ATTGGTGGTGGCTGGCAACTGGCATGGAATGGACTGATAAA---CGTGAGGATGGAAA 1470
QY 492 LysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArg 511
Db 1471 CAACAAGGAGGGTTTAAAGGATTATTATTACATGAGGAGGGAGTTTCTGTCATCTCGTCGT 1530
QY 512 GlySerIleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHis 531
Db 1531 GGATCCATTGTATCGATTCCCGGTGAAGGC-----GAATTTGTCCAG 1572
QY 532 AlaAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMet 551
Db 1573 GCTGCTGCCCTTGGTAAGCCAACCCGCTCTTACTCCAAGGAGCTTATTGATGGACACCCA 1632
QY 552 SerAspAlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAsp 571
Db 1633 GTTGGGCTTGCATGGTTTCAACCATCTGAGACAGCTTCAAAGGGGCCAAGTTGGAAGCT 1692
QY 572 LeuPheGluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGln 591
Db 1693 CTTCTTGAACCCAGGGTTAAGCATGCATGGTTGTTGGAGTTGGAATACAAATACTTCAG 1752
QY 592 GlnPheAlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGly 611
Db 1753 CAGTTTTTCAGGGATAAATGGGTTCTATATTACACACCTCAAATCCTTGAAGAGCCGGT 1812
QY 612 ValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSer 631
Db 1813 GTTGAAGTTCTTCTTTCAGATATAGGCATTGGCTCAGAGTCGGCATCATCTCTTATCAGT 1872
QY 632 SerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeu 651
Db 1873 GCTTTTCAACACCTTCTTGATGCTTCCCTGTATAGCGGTAGCCATGAAGCTCATGGATGT 1932
QY 652 SerGlyArgArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIle 671
Db 1933 TCAGGCAGAAGGCAGTTGCTACTTACTACAAATCCCCGCTGCTGATTGTGTCTCATTTATT 1992
QY 672 LeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrVal 691
Db 1993 TTGGTCATTGGAAGCCTGGTAAATTTTGGCAATGTGCGCCATGCAGCAATCTCAACAGTA 2052
QY 692 SerValIleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeu 711
Db 2053 TGGCTTGTGGTTTATTCTGCTGCTTTGTGATGGGTATGGACCAATTCCAAACATCCTT 2112
QY 712 CysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThr 731
Db 2113 TGCTCAGAGATTTTCCCACTAGGTGCGTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2172
QY 732 PheTrpIleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGly 751
Db 2173 TTCTGGATTGGAGACATCATCATCATCTACTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2232
QY 752 LeuAlaGlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPhe 771
Db 2233 CTTGGTGGTGTATTCCGCCATTACGCAGTTGTTTGTTCATCTCTCGGCTCTTTAGGA 2292
QY 772 LeuLysValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaVal 791
Db 2293 TTGAAGGTTCCAGAAACAAAGGCATGCCCTTGAAGTCATCTCTGAATTTCTTTCTGTT 2352
QY 792 GlyAlaLysGlnAlaAlaAla 798
Db 2353 GGAGCAAGCAGGCTGCTTCT 2373

Db 1240 GGATGGCAAGTGGCATGGAATGGACGGAAGAGAGATGAATCGGGACAGAAAGAA--- 1296

Qy 495 GlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgGlySerIle 514

Db 1297 -----GAAGGTTCCAGGATCTCGACGTGGCTCAATT 1329

Qy 515 ValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaAla 534

Db 1330 GTTTTCATTGCCTGGTGGTGATGGAACCGGTGAG--GCAGATTGTGACAAGCGTCTGCT 1386

Qy 535 LeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAspAla 554

Db 1387 TTGGTTAGCCAACAGACTCTTTATTCCAAAGACCTTCTCAAAGAACATACAATTGGTCTCT 1446

Qy 555 AlaMetValHisProSerGluValAlaAlaLysGlySerArgTyrTrpLysAspLeuPheGlu 574

Db 1447 GCTATGGTACATCCATCCGAA--ACAACATAAAGGGTCAATTGGGATGATCTTCATGAT 1503

Qy 575 ProGlyValArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnGlnPheAla 594

Db 1504 CCTGGAGTCAAGCGTGCATTAGTCGTAGGAGTTGGACTTCAAATACTTCAGCAGTTCTCA 1563

Qy 595 GlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyValAlaVal 614

Db 1564 GGCATCAACGGAGTTCTTTACTACACACCGCAAAATCCTTGAGCAGCGGGTGTCCGGATC 1623

Qy 615 IleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThr 634

Db 1624 CTACTATCGAACATGGGGATTAGTTCTTCTCAGCATCCTTACTTATAAGTGCATTGACA 1683

Qy 635 ThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArg 654

Db 1684 ACCTTTGTGATGTACTGCAATAGCTGTTGCAATGAGGCTCATGGATCTTTCTGGTCGA 1743

Qy 655 ArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuValVal 674

Db 1744 AGGACCTTGCTTCTCACCACGATACCAATCCTGATAGCATCTCTATTGGTTTAGTAATC 1803

Qy 675 SerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIle 694

Db 1804 TCAAATCTTGTTACATGAACAGCATTTGTCAGCGGGTCTTATCAACCGTAAGCGTTGTG 1863

Qy 695 ValTyrPheCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAlaGlu 714

Db 1864 CTCTACTTCTGCTTCTCGTGATGGGTTTCGGTCTCTGCTCCAAACATCCTCTGTTCAAG 1923

Qy 715 IlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIle 734

Db 1924 ATTTTCCAACTCGAGTCCGGGAATCTGCATCGCCATCTCGGCACCTTCTGGATC 1983

Qy 735 GlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGly 754

Db 1984 TGTGACATAATCGTCACTTACAGTCTCCCCGCTGCTGCTCAAATCCATGGACTAGCTGGT 2043

Qy 755 ValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeuLysVal 774

Db 2044 GTGTTTGGAAATGTACGCAATCGTATGTTGCAATTTCATGGGTCTTTTGTTTCAATAAGTC 2103

Qy 775 ProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAlaLys 794

Db 2104 CCGGAAACTAAAGGCATGCCACTTGAAGTCATCACAGAGTTCTTTTCTGTTGGAGCTAGA 2163

Qy 795 GlnAlaAlaAla 798

Db 2164 CAAGCTGAAGCT 2175

RESULT 10

ABZ14449

ID ABZ14449 standard; DNA; 2205 BP.

XX ABZ14449;

AC ABZ14449;

XX 21-JAN-2003 (first entry)

DT

XX Arabidopsis thaliana stress regulated gene SEQ ID NO 2254.

DE Arabidopsis thaliana; plant; gene; stress; transgenic; ds.

XX Arabidopsis thaliana.

OS Arabidopsis thaliana.

XX WO200216655-A2.

PN 28-FEB-2002.

PD 24-AUG-2001; 2001WO-US026685.

XX 24-AUG-2000; 2000US-0227866P.

PR 26-JAN-2001; 2001US-0264647P.

PR 22-JUN-2001; 2001US-0300111P.

XX (SCRI) SCRIPPS RES INST.

PA (SYGN) SYNGENTA PARTICIPATIONS AG.

XX Harper JF, Kreps J, Wang X, Zhu T;

PI WPI; 2002-304127/34.

XX Identifying a stress condition to which a plant cell has been exposed and producing plants with increased tolerance to these abiotic stresses.

PT Claim 144; SEQ ID NO 2254; 577pp + Sequence Listing; English.

XX The invention relates to identifying a stress condition to which a plant cell has been exposed, comprising: (a) contacting nucleic acid representative of expressed polynucleotides in the plant cell with an array or probes representative of the plant cell genome; and (b) detecting a profile of expressed polynucleotides in the plant cell characteristic of a stress response. The method is useful in the production of transgenic plants, cells and seeds and in producing plants with increased tolerance to abiotic stress. The present sequence is that of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used in methods of the invention. Note: The sequence data for this patent is not represented in the printed specification but is based on sequence information supplied to Derwent by the European Patent Office

XX Sequence 2205 BP; 521 A; 524 C; 521 G; 639 T; 0 U; 0 Other;

Qy Alignment Scores:

Pred. No.: 2.88e-189 Length: 2205

Score: 2275.00 Matches: 462

Percent Similarity: 75.89% Conservative: 111

Best Local Similarity: 61.19% Mismatches: 146

Query Match: 55.83% Indels: 36

DB: 6 Gaps: 14

US-10-051-909-32 (1-800) x ABZ14449 (1-2205)

Qy 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrp 75

Db 1 ATGAAGGGAGCGACTCTCGTTGCTCTCGCCGCCACAAATCGGCAATTTCTTACAAGGATGG 60

Qy 76 AspAsnAlaThrIleAlaAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsn 95

Db 61 GACAATGCCACCATTGCTGGAGCTATGGTTTATATCAACAAGACTTGAATCTA----- 114

Qy 96 GluProThr---ValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleVal 114

Db 115 ---CCAACTCTGTTCAAGGCTTGTGCTGTTGCTATGTCATGTCGGTGCAACGGTCAATC 171

Qy 115 ThrThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSer 134

Db 172 ACGACTTGCTCAGGACCGATATCTGATTGGCTCGGACAGACGCCCATGCTCATTTATCA 231

Qy 135 SerIleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyrValLeu 154

Db 232 TCAGTTATGTATTTCGTCTGCGGTTTGTGATAAAGTTGTGGTCTCCCAATGTCATGTTCTG 291

QY	155	LeuLeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeu	174
DB	292	TGCTTTGCTAGGCTCTTAATGGGTTTGGTGGCGGCTCGCGGTACACTTGTCCCTGTT	351
QY	175	TyrIleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPhe	194
DB	352	TACATTTCTGAAACCGCTCCTCCGGAGATCAGAGACAGTTAAATACTCTCCCTCAGTTT	411
QY	195	SerGlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerPro	214
DB	412	CTTGGCTCTGGTGAATGTTTTTGTCACTACTGTATGGTTTTTCACTATGTCCCTGAGTGAC	471
QY	215	SerProAspTirArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGly	234
DB	472	TCCCTAGCTGGAGAGCCATGCTCGGTGCTCTCGATCCCTTCTCTCTTTATTGTTT	531
QY	235	LeuThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMetAla	254
DB	532	CTCACGGTGTITTTATTTGCCCGAGTCTCCTCGTGGCTGGTTAGTAAAGGAAGATGGAC	591
QY	255	GluAlaLysLysValLeuGlnLysLeuArgGlyLysAspValSerGlyGluLeuSer	274
DB	592	GAGGCTAAGCGAGTCTTCAACAGTTATGTGGCAGAGAAAGATGTTACCGATGAGATGGCT	651
QY	275	LeuLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleGly	294
DB	652	TTACTAGTTGAGGACTAGATATAGGAGGAGAAAAACAATGGAAGATCTCTAGTAACT	711
QY	295	ProAlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGlu---	313
DB	712	TTGGAGGATCATGAAGGTGATGATACACTTGAACCGTTGATGAGGATGGACAAATGCGG	771
QY	314	LeuTyrGlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMet	333
DB	772	CTTTATGGAAACCCACGAGAAATCAATCGTACCTTGCTAGACCTGTCCCAGAACAA	825
QY	334	LeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValPro	353
DB	826	---AATAGCTCACTTGGGCTACGCTCTCGCCACGGAAGCTTAGCAAAACCAAGCATGATC	882
QY	354	LeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly	373
DB	883	CTTAAAGATCCGCTCGTCAATCTTTTGGCAGTCTCCACGAGAAGATGCCAGAAAGCAGGC	942
QY	374	GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln	393
DB	943	GGAACACTCGGAGTGGGATTTTCCCTCATTTCCGAAGCATGTTTCAGTACTACTGCCGAT	1002
QY	394	-----HisAlaLysAsnGluGlnTrpAsp-----GluGluAsnLeuHisArgAsp	408
DB	1003	CGCCTCACGGTAAACCGGCTCATTTGGGAAAAGGACATAGAGAGCCATTACAAACAAAGAC	1062
QY	409	AspGluGluTyrAlaSerAspGlyAlaGlyGlyAspTyrGluAsp-----AsnLeu	425
DB	1063	AATGATGACTATGCGACTGATGATGGTGGGGTGATGATGATGATCGGACACCATTTG	1122
QY	426	HisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIleValHisHis	445
DB	1123	CGTAGCCCCCTTAATGTGCGGCCAGACCACAAAGCATGGAC---AAGGATATGATGCCACAT	1179
QY	446	GlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeuGlyGluGlyAsp	465
DB	1180	CCTACAAAGTGAAGCACTTTAAGCATGAGACGACACAGTACGCTTATGCAA--GGCAAC	1236
QY	466	GlyValSerSerThrAspIleGlyGlyTrpGlnLeuAlaTirLysTrpSerGluLys	485
DB	1237	GGCGAAAGTAGCATGGGAATTTGGTGGTGGTGGCATATGGGATATAGATACGAAAAACGAT	1296
QY	486	GluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGly	505
DB	1297	GAA-----TACAAGAGGTATTATCTTAAAGAAAGATGGA	1329

QY	506	ValProGlySerArgArgGlySerIleValSerLeuProGlyGlyGlyAspValPheGlu	525
DB	1330	GCTGAA--TCTCGCCGTGGCTCGATCATCTCTATTCCCAGGATCCGGAT-----GGT	1380
QY	526	GlySerGluPheValHisAlaAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGly	545
DB	1381	GGAGGCAGCTACATTACGCTTCTGCCCTTGTAAGCAGATCTGTTCTTGGTCCTAAATCA	1440
QY	546	LeuAlaGluProArgMetSerAspAlaAlaMetValHisProSerGluValAlaAlaLys	565
DB	1441	-----GTTTCATGGATCCGCCATGGTCCCCCGAGAAAATAATGCTGCCTCT	1485
QY	566	GlySerArgTrpLysAspLeuPheGluProGlyValArgArgAlaLeuLeuValGlyVal	585
DB	1486	GGACCACCTCTGGTCTGCTCTTCTTGAACCTGGTGTAAAGCGTCCCTTGGTGTGGTGTGTC	1545
QY	586	GlyIleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyValLeuTyrtyrThrProGln	605
DB	1546	GGCATTCAAATACTGCAGCAGTTTTTCAGGTATCAATGGAGTTCTCTACTACACTCCTCAG	1605
QY	606	IleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSer	625
DB	1606	ATTCTCGAACGGGCTGGCGTAGATATTCTTTCGAGCCTCGGACTAAGTTCCATCTCT	1665
QY	626	AlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAla	645
DB	1666	CGTCACTTCCTCATCAGCGGTTTAAACAACATTACTCATGCTCCAGCCATTGTCGTTGCC	1725
QY	646	MetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuLeuGlyThrlleProileLeu	665
DB	1726	ATGAGACTCATGGATGTATCCGGAAGAAGTTCATTACTTCTCTGGACAATCCAGTTCTC	1785
QY	666	IleAlaSerLeuValIleLeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHis	685
DB	1786	ATTGTCTCACTGTGTCCTTGTTCATCAGCAGCTCATCCACATCAGCAAGTCTGTGAAC	1845
QY	686	AlaLeuLeuSerThrValSerValIleValTyrPheCysCysPheValMetGlyPheGly	705
DB	1846	GCAGCACTCCACAGGTTGTGTCGTCTACTTCTGCTTCTTCGTGATGGGTACGGT	1905
QY	706	ProIleProAsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIle	725
DB	1906	CCCATTCCAAACATCCTCTGTTCTGAAATCTTCCAACAAGAGTCCGTGGTCTCTGCATC	1965
QY	726	AlaIleCysAlaPheThrPheTrpIleGlyAspIleIleValThrTyrSerLeuProVal	745
DB	1966	GCCATATGTGTATGGTCTTTTGGATTGGAGACATTATTGTACGTACTCACTTCCCGTT	2025
QY	746	MetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeuIle	765
DB	2026	CCTCCTCAGCTCGATCGGACTAGTTGGTGTGTTTCAGCATTTACGCTGCGGTTTGCATTATC	2085
QY	766	SerPheValPheValPheLeuLysValProGluThrLysGlyMetProLeuGluValIle	785
DB	2086	TCAATGGATCTCGTTTACATGAAGTCCCGAGACTAAGGCATGCCCTTTGGAAGTTATC	2145
QY	786	ThrGluPhePheAlaValGlyAlaLysGlnAlaAlaLysAla	800
DB	2146	ACAGACTACTTGCCTTTGGAGCT---CAAGCTCAAGCTTCTGCT	2187
RESULT 11			
ID	ADG87978 standard; cDNA; 2205 BP.		
XX	ADG87978;		
XX			
DT	22-APR-2004 (first entry)		
DE	A. thaliana RPP4-upregulated pathogen infection-related gene #420.		
XX	Pathogen infection-related gene; plant; Peronospora parasitica;		
KW	defence mechanism; RPP4; pathogen resistance; transgenic plant; oomycete;		
KW	fungus; bacterium; virus; nematode; insect; aphid; gene; ss		

```
XX Arabidopsis thaliana.
XX WO200222675-A2.
XX 21-MAR-2002.
XX 14-SEP-2001; 2001WO-US028506.
XX 15-SEP-2000; 2000US-0232778P.
XX 22-JUN-2001; 2001US-0300183P.
XX (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX (UYN-) UNIV NORTH CAROLINA.
XX (GLAZ/) GLAZEBROOK J.
XX (WANG/) WANG X.
XX (DANG/) DANGL J L.
XX (EULG/) EULGEM T.
XX (ZHUT/) ZHU T.
XX Glazebrook J, Wang X, Dangl JL, Eulgem T, Zhu T;
XX WPI; 2002-292409/33.
XX Novel isolated polynucleotide, useful for conveying pathogen resistance
XX to plants, and for identifying plants infected with a pathogen.
XX Claim 3; SEQ ID NO 420; 605pp; English.
XX The invention relates to 691 Arabidopsis thaliana genes (ADG87559--
XX ADG87557) whose expression is altered in response to pathogen infection,
XX and to homologues of these genes from other plants or fungi, especially
XX from maize, soybean, barley, alfalfa, sunflower, canola (oilseed rape),
XX cotton, peanut, sorghum, tobacco, sugarbeet, rice or wheat. The
XX expression of genes of the invention was upregulated or downregulated in
XX Arabidopsis plants infected with the oomycete Peronospora parasitica,
XX indicating that they play a role in defence mechanisms. The genes of the
XX invention are regulated by RPP7 or RPP8 which act via unconventional
XX signalling cascades, or by the RPP4-dependent pathway. The invention also
XX relates to polypeptides encoded by the pathogen infection-related genes;
XX promoter motifs from pathogen infection-related genes (ADG88243-ADG88327)
XX ; expression cassettes, host cells and pathogen-resistant transgenic
XX plants and their progeny comprising a polynucleotide of the invention;
XX and a method of identifying a plant cell infected with a pathogen. The
XX polynucleotide sequences and methods of the invention are useful for
XX identifying plants infected with a pathogen, and for conferring
XX resistance to pathogens such as oomycetes, fungi, bacteria, viruses,
XX nematodes and insects (e.g., aphids). The present sequence represents an
XX Arabidopsis thaliana gene whose expression is altered in response to
XX Peronospora parasitica infection. Note: The sequence data for this patent
XX did not form part of the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 2205 BP; 521 A; 524 C; 521 G; 639 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 2.88e-189 Length: 2205
XX Score: 2275.00 Matches: 462
XX Percent Similarity: 75.89% Conservative: 111
XX Best Local Similarity: 61.19% Mismatches: 146
XX Query Match: 55.83% Indels: 36
XX DB: 6 Gaps: 14
XX
XX US-10-051-909-32 (1-800) x ADG87978 (1-2205)
XX
XX 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrp 75
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 1 ATGAAGGAGCGACTCTCGTTGCTCTCGCGCCACAAATCGGCAATTCTTACAAGGATGG 60
XX
XX 76 AspAsnAlaThrIleAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsn 95
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 61 GACAATGCCACCAATTGCTGGAGCTATGTTTATATCAACAAGACTTGAATCTA----- 114
```

```
Qy 96 GluProThr---ValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleVal 114
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 115 ---CCAACCTCTGTTCAAGGTCTTGTGCTGCTATGTCATTCGTCGTCACGGTCATC 171
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 115 ThrThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSer 134
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 172 ACGACTTGCTCAGGACCGATATCTGATTGGCTCGGCAGACGCCCATGCTCATTTATCA 231
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 135 SerIleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyrValLeu 154
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 232 TCAGTTATGATTTCGTCGCGGTTTGATAATGTTGTGGTCTCCCAATGCTATGTTCTG 291
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 155 LeuLeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeu 174
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 292 TGCTTTGCTAGGCTTCTTAATGGGTTTGGTGCCGGGCTCGCGGTTACACTTGTCCTGTT 351
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 175 TyrIleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPhe 194
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 352 TACATTTCTGAACCCGCTCCTCCGGAGATCAGAGACAGTTAAATACTCTCCCTCAGTTT 411
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 195 SerGlySerGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerPro 214
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 412 CTGGCTCTGGTGAATGTTTGTCTACTGTCATGCTGTTTCACTATGTCCTCAGTGAC 471
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 215 SerProAspTyrArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGly 234
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 472 TCCCTAGCTGGAGAGCCATGCTCGGTGTCCTCTCGATCCCTTCTCTTCTTATTGTTT 531
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 235 LeuThrIlePheTyrLeuProGluSerProArgTyrLeuValSerLysGlyArgMetAla 254
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 532 CTCACGGTGTATTTTCCCGAGTCTCCTCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 591
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 255 GluAlaLysLysValLeuGlnLysLeuArgGlyLysAspValSerGlyGluLeuSer 274
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 592 GAGGTAAGCGAGTTCTCAACAGTTATGTGGCAGAGAGAGATGTTACCGATGAGATGGCT 651
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 275 LeuLeuLeuGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIleGly 294
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 652 TTAAGTTGAAGGACTAGATATAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 711
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 295 ProAlaThrGluAlaAlaAspLeuValThrAspGlyAspLysGlu---GlnIleThr 313
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 712 TTGGAGGATCATGAAGGTGATGATACACTTGAAACCGTTGATGAGGATGGACAAATGCGG 771
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 314 LeuTyrGlyProGluGlyGlnSerTyrPheAlaArgProSerLysGlyProIleMet 333
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 772 CTTTATGGAAACCCACGAGATCAATCGTACCTGCTAGACCTGTCCCAGAACAA----- 825
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 334 LeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValPro 353
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 826 ---AATAGCTCACTTGGGCTACGCTCTCGCCACGGAAGCTTAGCAAAACCAAGCATGATC 882
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 354 LeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly 373
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 883 CTTAAAGATCCGCTCGTCAATCTTTTGGCAGTCTCCACGAGAGATGCCAGAACGAGGC 942
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 943 GGAAACACTCGGAGTGGGATTTTCCCTCATTTCCGGAAGCATGTTTCAGTACTACTCCGAT 1002
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 394 -----HisAlaLysAsnGluGlnTyrAsp-----GluGluAsnLeuHisArgAsp 408
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1003 CGGCCTCACGGTAAACCGGCTCATTTGGGAAAGGACATAGAGAGCCATTACACAAAGAC 1062
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 409 AspGluGluTyrAlaSerAspGlyAlaGlyGlyAspTyrGluAsp-----AsnLeu 425
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1063 AATGATGACTATCGGACTGATGATGGTGGGCGGTGATGATGATGATGATGATGATGATG 1122
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 426 HisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIleValHisHis 445
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1123 CGTAGCCCCCTTAATGTGCGGCCAGACCAACAGCATGGAC---AAGGATATGATCCCACAT 1179
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```


QY 446 GlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeuGlyGluGlyGlyAsp 465
DB 1180 CCTACAACTGGAAGCACTTTTACCATGAGACGACACAGTACGCTTATGCAA---GGCAAC 1236
QY 466 GlyValSerSerThrAspIleGlyGlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLys 485
DB 1237 GCGAAAGTAGCATGGAAATTGGTGGTGGTGGCATATCGGATATAGATACGAAACGAT 1296
QY 486 GluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGly 505
DB 1297 GAA-----TACAAGAGGTATTATCTTAAAGAAAGATGGA 1329
QY 506 ValProGlySerArgArgGlySerIleValSerLeuProGlyGlyGlyAspValPheGlu 525
DB 1330 GCTGAA---TCTCGCGTGGCTCGATCATCTCTATTCCCGAGGTCCGGAT-----GGT 1380
QY 526 GlySerGluPheValHisAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGly 545
DB 1381 GGAGCGAGCTACATTACGCTTCTGCTTCTTGAACCTGGTGTAAAGCAGATCTGTCTTGGTCTTAATCA 1440
QY 546 LeuAlaGluProArgMetSerAspAlaAlaMetValHisProSerGluValAlaAlaLys 565
DB 1441 -----GTTTCATGGATCCGCCATGGTTCCTCCCGGAGAAATTCCTGCTCT 1485
QY 566 GlySerArgTrpLysAspLeuPheGluProGlyValArgAlaLeuValGlyVal 585
DB 1486 GGACCACCTGGTCTGCTCTTCTTGAACCTGGTGTAAAGCAGTCTGTCTTGGTGTGGTGTCT 1545
QY 586 GlyIleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyValLeuTyrTrpProGln 605
DB 1546 GGCATTCAATACTGCAGCAGTTTTCAGGTATCAATGGAGTCTCTACTACATCTCTCAG 1605
QY 606 IleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSer 625
DB 1606 ATTCTCGAACGGCTGGCGTAGATATTCTTCTTCGAGCCTCGGACTAAGTTCCATCTCT 1665
QY 626 AlaSerIleLeuSerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAla 645
DB 1666 GCGTCATCTCATCAGCGGTTTAAACAACATTACTCATGTCCAGCCATTGTCTGCTGCC 1725
QY 646 MetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuLeuGlyThrIleProIleLeu 665
DB 1726 ATGAGACTCATGGATGTATCCGGAAGAAGGTCAATCTCTCTGGACAATCCCACTTCTC 1785
QY 666 IleAlaSerLeuValIleLeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHis 685
DB 1786 ATTGTCTCACTTGTCTGCTCTTGTATCAGCAGCTCATCCACATCAGCAAAAGTCTGAAC 1845
QY 686 AlaLeuLeuSerThrValSerValIleValTyrPheCysCysPheValMetGlyPheGly 705
DB 1846 GCAGCACTCTCCACAGGTTGTGTCTGCTCTACTTCTGCTTCTGTTGATGGGTACGGT 1905
QY 706 ProIleProAsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIle 725
DB 1906 CCCATTCCAAACATCCTCTGTCTGAAATCTTCCCAACAAGAGTCCGTGCTCTCTGATC 1965
QY 726 AlaIleCysAlaPheThrPheTrpIleGlyAspIleIleValThrTyrSerLeuProVal 745
DB 1966 GCCATATGTGTATGGTCTTTTGGATTGGAGACATTATTGTACAGTACTCACTTCCCGTT 2025
QY 746 MetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeuIle 765
DB 2026 CTCCTCAGCTCGATCGGACTAGTTGGTGTGTTTTCAGCATTTACGCTGCGGTTTGGTTATC 2085
QY 766 SerPheValPheValPheLeuLysValProGluThrLysGlyMetProLeuGluValIle 785
DB 2086 TCATGGATCTCGTTTACATGAAGTCCCGGAGACTAAAGCATGCCCTTTGGAAGTTATC 2145
QY 786 ThrGluPhePheAlaValGlyAlaLysGlnAlaAlaLysAla 800
DB 2146 ACAGACTACTTTCCTTTGGAGCT---CAAGCTCAAGCTTCTGCT 2187

ABK51966
ID ABK51966 standard; cDNA; 1692 BP.
XX
AC ABK51966;
XX 27-AUG-2002 (first entry)
XX Soybean cDNA clone ss1.pk0022.f1 encoding sugar transport protein.
XX Soybean; Arabidopsis thaliana-like sugar transport protein;
KW carbohydrate transport; grain filling; annual field crop; plant;
KW clone ss1.pk0022.f1; gene; ss.
XX Glycine max.
OS
XX Key Location/Qualifiers
CDS 9..1469
FT /*tag= a
FT /partial
FT /product= "Portion of a soybean Arabidopsis thaliana-like
FT sugar transport protein"
FT /note= "This sequence lacks a start codon"
XX
PN US6383776-B1.
XX
PD 07-MAY-2002.
XX
PF 14-APR-1999; 99US-00291922.
XX
PR 24-APR-1998; 98US-0083044P.
XX (DUPO) DU PONT DE NEMOURS & CO E I.
XX
XX Allen SM, Hitz WD, Kinney AJ, Tingey SV;
PI WPI; 2002-453364/48.
XX P-PSDB; AAU97205.
XX New nucleic acid encoding plant sugar-transport proteins, useful for
PT preparing transgenic plants with altered carbohydrate distribution.
XX Claim 3; Col 41-44; 54pp; English.
XX The present invention relates to the isolation of plant polynucleotide
CC sequences encoding an Arabidopsis thaliana-like sugar transport protein
CC or Beta vulgaris-like sugar transport protein. The polynucleotide
CC sequences are useful for altering the level of sugar transport proteins
CC in plants, i.e. for control of carbohydrate transport and distribution in
CC plant cells, e.g. during grain filling of annual field crops (e.g. corn,
CC rice, soybeans, and wheat), and, for studying carbohydrate flows and
CC sugar transport. The polynucleotide sequences can also be used to isolate
CC cDNA sequences and genes that encode homologues of the new proteins. The
CC present sequence encodes a portion of a soybean Arabidopsis thaliana-like
XX sugar transport protein
SQ Sequence 1692 BP; 430 A; 339 C; 400 G; 523 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2.19e-136 Length: 1692
Score: 1670.50 Matches: 333
Percent Similarity: 79.75% Conservative: 57
Best Local Similarity: 68.10% Mismatches: 84
Query Match: 40.99% Indels: 15
DB: 6 Gaps: 8

US-10-051-909-32 (1-800) x ABK51966 (1-1692)

QY 308 AspLysGluGlnIleThrLeuTyrGlyProGluGluGlyGlnSerTrpIleAlaArgPro 327
DB 21 GAAAAGATCAATTAAGTTGTATGGACCAGAACAGGCCAGTCTGGTGTGCTAGACCT 80
QY 328 SerLysGlyProIleMetLeuGlySerValLeuSerLeuAlaSerArgHisGlySerMet 347

Db 81 GTTGCTGGACCAAAATCTGTTGGC-----CTTGTATCTAGGAAAGGAGCATG 128

QY 348 ValAsnGlnSerValProLeuMetAspProIleValThrLeuPheGlySerValHisGlu 367

Db 129 GCAAATCCAAGC---AGTCTAGTGGACCTCTAGTGACCCTCTTTGGTAGTGATCATGAG 185

QY 368 AsnMetProGlnAlaGlyGlySerMetArgSerThrLeuPheProAsnPheGlySerMet 387

Db 186 AAGCTCCAGAAACAGGA-----AGCACCTTTTCCACACTTTGGGAGTATG 233

QY 388 PheSerValThrAspGlnHisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArg 407

Db 234 TTCAGTGTTGGGGAAATCAGCCCAAGGAATGAAGATTGGGATGAGGAAAGCCTAGCCAGA 293

QY 408 AspAspGluCluTyrAlaSerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSer 427

Db 294 GAGGGTGATGATTATGCTCTGAT-----GCTGGTGATTCTGATGACAAATTTGCAGAGT 347

QY 428 ProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHis 447

Db 348 CCATTGATCTACGTCACGTCAAACAACGAGTCTGGAT---AAGGACATACCTCTCATGCCCAT 404

QY 448 ArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeu---GlyGluGlyGlyAspGly 466

Db 405 AGTAACCTTGCA---AGCATGAGGCAAGGTAGTCTTTTACATGGAAATTCAGGAGAACCC 461

QY 467 ValSerSerThrAspIleGlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGlu 486

Db 462 ACTGGTAGTACTGGGATTGGTGGTGGTGGCAGCTAGCATGAAATGGTCTGAAAGAGAG 521

QY 487 GlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGlyVal 506

Db 522 GGCCCGAGATGGAAGAAGGAGGTGGCTTCAAGAGAAATATATTACACCAAGATGGTGGT 581

QY 507 ProGlySerArgArgGlySerIleValSerLeuProGlyGlyGlyAspValPheGluGly 526

Db 582 TCTGGATCTAGACGTGGGTCTGTGGTTTCACTCCCT---GGCGTGATTTACCAACTGAC 638

QY 527 SerGluPheValHisAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeu 546

Db 639 AGTGAGGTGTACAGGCTGCTGCTCTGGTGAGTCAGCCTGCCCTTTATATAGGAGACCTT 698

QY 547 AlaGluProArgMetSerAspAlaAlaMetValHisProSerGluValAlaAlaLysGly 566

Db 699 ATGCGTCAACGCCAGCTGGACCAGCTATGATTTCATCCCTCTGAAACAATTGCAAAAGGG 758

QY 567 SerArgTrpLysAspLeuPheGluProGlyValArgArgAlaLeuValGlyValGly 586

Db 759 CCAAGTTGGAGTGATCTTTTGAACCTGGGGTGAAGCATGCATGATTGTGGGGTGGGA 818

QY 587 IleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyValLeuTyrThrProGlnIle 606

Db 819 ATGCAAAATCTTCAGCAGTCTCTCTGGTATAAATGGGTCTCTACTATACGCCTCAAAAT 878

QY 607 LeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAla 626

Db 879 CTTGAGCAGGAGGTGTTGGTTATCTTCTTTCAAGCCTAGGCCTGGTTCCTACTTCTTCA 938

QY 627 SerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMet 646

Db 939 TCCTTTCTTATTAGTGGGTGACAACTTGTTGATGCTTCTCTGTATAGCCATTGGCCATG 998

QY 647 LeuLeuMetAspLeuSerGlyArgArgPheLeuLeuLeuGlyThrIleProIleLeuIle 666

Db 999 AGGCTCATGGATATTTCAGGCAGAGGACTTGTGCTCAGTACAATCCCGTCCCTAATA 1058

QY 667 AlaSerLeuValIleLeuValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAla 686

Db 1059 GCAGCTCTTCTCATATTAGTCTCGGAAGCTCTGTGGATTGGGATCCCACTGCAAAATGCA 1118

QY 687 LeuLeuSerThrValSerValIleValTyrPheCysCysPheValMetGlyPheGlyPro 706

Db 1119 TCAATCTCAACCATTAGTGTATTGTTCTATTCTGTTTCTTCTGATGGGATTTGGACCA 1178

QY 707 IleProAsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIleAla 726

Db 1179 ATTCCTAATATACTTTGTGCAGAGATCTTCCCCACTCGAGTTCGIGGTCTCTGCATTGCT 1238

QY 727 IleCysAlaPheThrPheTrpIleGlyAspIleIleValThrTyrSerLeuProValMet 746

Db 1239 ATTTGTGCCCTTACCTTTTGGATCTGTGATATCATTTGCCTACACTCCCAGTTATG 1298

QY 747 LeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeuIleSer 766

Db 1299 CTCAATTCTGTAGGCCTCGTGGTGTGTTTGGTATTTATGCTGCTGCTCATAGCA 1358

QY 767 PheValPheValPheLeuLysValProGluThrLysGlyMetProLeuGluValIleThr 786

Db 1359 TGGGTGTTTGTCTTTTGAAGTTCCAGAAACCAAGGCGCATGCCACTGGAAGTGATCATT 1418

QY 787 GluPhePheAlaValGlyAlaLysGln 795

Db 1419 GAGTTCTTCTCTCGGAGCAAAACAG 1445

RESULT 13

ABX93202

ID ABX93202 standard; cDNA; 1692 BP.

XX

AC ABX93202;

XX

DT 29-MAY-2003 (first entry)

XX

DE cDNA encoding soybean sugar transport protein #2.

XX

KW Arabidopsis thaliana-like sugar transport protein; corn; rice; wheat;

KW plant sugar transport protein; carbohydrate transport; soybean;

KW carbohydrate distribution; plant; gene; ss.

XX

OS Glycine max.

OS

PN US2002178468-A1.

XX

PD 28-NOV-2002.

XX

PF 17-JAN-2002; 2002US-00051902.

XX

PR 24-APR-1998; 98US-0083044P.

PR 14-APR-1999; 99US-00291922.

XX

PA (ALLE/) ALLEN S M.

PA (HITZ/) HITZ W D.

PA (KINN/) KINNEY A J.

PA (TING/) TINGEY S V.

XX

Allen SM, Hitz WD, Kinney AJ, Tingey SV;

WPI; 2003-340957/32.

P-PSDB; ABU08330.

PT Novel plant sugar transport proteins and nucleic acid encoding the

PT protein useful for producing transgenic plants having altered levels of

PT sugar transport protein.

XX

PS Claim 2; Page 22-23; 56pp; English.

XX

The present invention relates to the isolation of Arabidopsis thaliana-
like or Beta vulgaris-like sugar transport proteins, and the
polynucleotide sequences encoding them. The plant sugar transport
proteins of the invention have been isolated from corn, rice, soybean,
and wheat. The polypeptides of the invention may be used for altering the
level of expression of a sugar transport protein in a host cell, by
transforming a host cell with a chimeric construct encoding all, or a
portion of the sugar transport protein, in sense or antisense
orientation. Particularly, the polypeptides may provide a means to
control carbohydrate transport and distribution in plants. ABX93198-
ABX93205 represent cDNA sequences encoding Arabidopsis thaliana-like

CC sugar transport proteins

XX Sequence 1692 BP; 430 A; 339 C; 400 G; 523 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2.19e-136 Length: 1692
Score: 1670.50 Matches: 333
Percent Similarity: 79.75% Conservative: 57
Best Local Similarity: 68.10% Mismatches: 84
Query Match: 40.99% Indels: 15
DB: 8 Gaps: 8

US-10-051-909-32 (1-800) x ABX93202 (1-1692)

Qy 308 AsplysGluGlnIleThrLeuTyrGlyProGluGluGlyGlnSerTrpIleAlaArgPro 327
Db 21 GAAAAGATCAAAATTAAGTTGTATGGACCAGAACAGGCCAGTCTCTGGTTGCTAGACCT 80
Qy 328 SerLysGlyProIleMetLeuGlySerValLeuSerLeuAlaSerArgHisGlySerMet 347
Db 81 GTTGCTGGACCAAAATCTGTGGC-----CTGTATCTAGGAAAGGAAGCATG 128
Qy 348 ValAsnGlnSerValProIleMetAspProIleValThrLeuPheGlySerValHisGlu 367
Db 129 GCAATCCAAGC--AGTCTAGTGGACCTCTAGTGACCTCTTTGGTAGGTACATGAG 185
Qy 368 AsnMetProGlnAlaGlyGlySerMetArgSerThrLeuPheProAsnPheGlySerMet 387
Db 186 AAGCTCCCAAGAACAGGA-----AGCACCTTTTCCACACTTTGGGAGTATG 233
Qy 388 PheSerValThrAspGlnHisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArg 407
Db 234 TTCACTGTTGGGGAAATCACCCCAAGGAATGAAGATGGGATGAGGAAGCTAGCCAGA 293
Qy 408 AspAspGluGluTyrAlaSerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSer 427
Db 294 GAGGTGATGATTATGCTCTGAT-----GCTGGTGTCTGATGACAAATTCGACAGT 347
Qy 428 ProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHis 447
Db 348 CCATTGATCTCACGTCAAAACACGAGTCTGGAT---AAGGACATACCTCCTCATGCCAT 404
Qy 448 ArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeu---GlyGluGlyGlyAspGly 466
Db 405 AGTAACCTTGCA---AGCATGAGGCAAGGTAGTCTTTACATGGAAATTCAGGAGAACCC 461
Qy 467 ValSerSerThrAspIleGlyGlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGlu 486
Db 462 ACTGTAGTACTGGGATTGGTGGTGGTGGCAGCTAGCATGGAATGGTCTGAAAGAGAG 521
Qy 487 GlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGlyVal 506
Db 522 GGCCAGATGGAAGAAGAGAGGTGGCTTCAAGAGAATATATTACACCAAGATGGTGGT 581
Qy 507 ProGlySerArgArgGlySerIleValSerLeuProGlyGlyGlyAspValPheGluGly 526
Db 582 TCTGATCTAGACGTGGTCTGTGGTTTCACTCCCT--GGCGGTGATTACCAACTGAC 638
Qy 527 SerGluPheValHisAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeu 546
Db 639 AGTGAGTTGTACAGGCTGCTGCTCTGGTGTAGTCAGCTGCCCTTTATAATGAGGACCTT 698
Qy 547 AlaGluProArgMetSerAspAlaAlaMetValHisProSerGluValAlaAlaLysGly 566
Db 699 ATGCGTCAACGCCAGTTGGACCACTATGATTCTACCTCTGAAACAAATGCAAAAGGG 758
Qy 567 SerArgTrpLysAspLeuPheGluProGlyValArgArgAlaLeuLeuValGlyValGly 586
Db 759 CCAAGTTGGAGTGATCTTTTGAACCTGGGTGAAGCATGCATTTGATTGTGGGGTGGGA 818
Qy 587 IleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIle 606
Db 819 ATGCAAAATCTTCAGCAGCTTCTCTGGTATAAATGGGGTCTCTACTATACGCTCAATT 878

Qy 607 LeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAla 626
Db 879 CTTGAGCAGGCAGGTGTTGGTTATCTTCTTTCAAGCTAGGCCTTGGTTCTACTTCTTCA 938
Qy 627 SerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMet 646
Db 939 TCCTTTCTTATTAGTGGGTGACAACTTGTGTATGCTTCTTGTATAGCCATTGCCATG 998
Qy 647 LeuLeuMetAspLeuSerGlyArgArgPheLeuLeuLeuGlyThrIleProIleLeuIle 666
Db 999 AGGCTCATGGATATTTTCAGCAGAGGACTTTTGTCTGCTCAGTACAATCCCGCTCTAATA 1058
Qy 667 AlaSerLeuValIleLeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAla 686
Db 1059 GCAGCTCTTCTCATATTAGTCTCTGGGAGTCTTGTGGATTGGGATCCACTGCAAAATGCA 1118
Qy 687 LeuLeuSerThrValSerValIleValTyrPheCysCysPheValMetGlyPheGlyPro 706
Db 1119 TCAATCTCAACCATTAGTGTATTGTCTATTTCTGTCTTCTTGTCTATGGGATTGGACCA 1178
Qy 707 IleProAsnIleLeuCysAlaGluIlePhePheProThrArgValArgGlyLeuCysIleAla 726
Db 1179 ATTCTAATATATCTTTGTGCAGAGATCTTCCCCACTCGAGTTCGTGGTCTCTGCATTGCT 1238
Qy 727 IleCysAlaPheThrPheTrpIleGlyAspIleIleValThrTyrSerLeuProValMet 746
Db 1239 ATTTGTGCCCTTACCTTTTGGATCTGTGATATCATTTGTACCTACACACTCCAGTTATG 1298
Qy 747 LeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeuIleSer 766
Db 1299 CTCATTTCTGTAGGCCCTCGCTGGTGTGTTTGGTATTTATGCTGTGTGTGCTCATAGCA 1358
Qy 767 PheValPheValPheLeuLysValProGluThrLysGlyMetProLeuGluValIleThr 786
Db 1359 TGGGTGTTTGTCTTTTGAAGATTCCAGAAACCAAGGCATGCCACTGGAAAGTATCATT 1418
Qy 787 GluPhePheAlaValGlyAlaLysGln 795
Db 1419 GAGTTCTTCTCTGTCGGAGCAAAACAG 1445
RESULT 14
ADG47913
ID ADG47913 standard; cDNA; 1692 BP.
XX
AC ADG47913;
XX
DT 11-MAR-2004 (first entry)
XX
DE Soybean Arabidopsis-like sugar transport protein cDNA.
XX
KW Arabidopsis-like sugar transport protein;
KW Beta-vulgaris-like sugar transport protein; transgenic; physical mapping;
KW Soybean; plant; gene; ss.
XX
OS Glycine max.
XX
FH Key Location/Qualifiers
FT CDS 9..1469
FT /*tag= a
FT /product= "Soybean Arabidopsis-like sugar transport
FT protein"
FT /note= "CDS does not include start codon"
FT /partial
XX
PN US2002199217-A1.
XX
PD 26-DEC-2002.
XX
PF 17-JAN-2002; 2002US-00051909.
XX
PR 24-APR-1998; 98US-0083044P.
PR 14-APR-1999; 99US-00291922.

XX (HELE/) HELENTJARIS T G.
XX PA Helentjaris TG;
XX PI
XX WPI; 2004-040967/04.
DR P-PSDB; ADG47914.
XX
PT New isolated polynucleotide encoding a polypeptide having sugar transport
PT protein activity, for producing a transformed plant and for use as probes
PT in physical mapping.
XX
PS Disclosure; SEQ ID NO 9; 71pp; English.
XX
CC The invention relates to Arabidopsis-like or Beta-vulgaris-like sugar
CC transport proteins and their corresponding nucleic acid sequences. The
CC sequences of the invention are useful to transform a cell. These are also
CC useful to produce a transgenic plant. Probes derived from sequences
CC encoding sugar transport protein may be used for physical mapping. The
CC present sequence is Soybean Arabidopsis-like sugar transport protein
CC cDNA.
XX
SQ Sequence 1692 BP; 430 A; 339 C; 400 G; 523 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2.19e-136 Length: 1692
Score: 1670.50 Matches: 333
Percent Similarity: 79.75% Conservatives: 57
Best Local Similarity: 68.10% Mismatches: 84
Query Match: 40.99% Indels: 15
DB: 12 Gaps: 8

US-10-051-909-32 (1-800) x ADG47913 (1-1692)

QY	308	AspLysGluGlnIleThrLeuTyrGlyProGluGluGlnSerTrpIleAlaArgPro	327
Db	21	GAAAAGATCAAAATTAAAGTTGTATGGACCAGAACAGGCCAGTCCTGGGTGCTAGACCT	80
QY	328	SerLysGlyProIleMetLeuGlySerValLeuSerLeuAlaSerArgHisGlySerMet	347
Db	81	GTTGCTGGACCAAAATTCGTGGC-----CTTGATCTAGGAAAGGAGCATG	128
QY	348	ValAsnGlnSerValProLeuMetAspProIleValThrLeuPheGlySerValHisGlu	367
Db	129	GCAAAATCCAAGC--AGTCTAGTGGACCTCTAGTGACCCCTCTTTGGTAGTGTACATGAG	185
QY	368	AsnMetProGlnAlaGlyGlySerMetArgSerThrLeuPheProAsnPheGlySerMet	387
Db	186	AAGCTCCAGAAACAGGA-----AGCACCTTTTCCACACTTTGGGAGTATG	233
QY	388	PheSerValThrAspGlnHisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArg	407
Db	234	TTCAGTGTGGGGGAAATCAGCCAAGGAATGAAGATTGGGATGAGGAAAGCCTAGCCAGA	293
QY	408	AspAspGluGluTyrAlaSerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSer	427
Db	294	GAGGGTGATGATTATGTCTCTGAT-----CCTGGTGATTCTGATGACAAATTGCAGAGT	347
QY	428	ProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHis	447
Db	348	CCATTGATCTCACGTCAAAACAACGAGTCTGGAT--AAGGACATACCTCCTCATGCCCAT	404
QY	448	ArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeu--GlyGluGlyGlyAspGly	466
Db	405	AGTAACCTTGCA--AGCATGAGGCAAGGTAGTCTTTTACATGGAAATTCAGAGAACCC	461
QY	467	ValSerSerThrAspIleGlyGlyTyrGlnLeuAlaTrpLysTrpSerGluLysGlu	486
Db	462	ACTGGTAGTACTGGGATTGGTGGTGGTGGCAGCTAGCATGGAAATGGTCTGAAAGAGAG	521
QY	487	GlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGlyVal	506
Db	522	GGCCACAGATGGAAAGCAAGGAAGGTGGCTTCAAGACAATATATTTACACCAAGATGGTGGT	581

KW ss.
XX Triticum aestivum.
OS
XX
FH Key Location/Qualifiers
FT CDS 3..1040
FT /*tag= a
FT /partial
FT /product= "Portion of a wheat Arabidopsis thaliana-like
FT sugar transport protein"
FT /note= "This sequence lacks a start codon"
XX
PN US6383776-B1.
XX
XX
PD 07-MAY-2002.
XX
XX 14-APR-1999; 99US-00291922.
XX
XX 24-APR-1998; 98US-0083044P.
XX
XX (DUPO) DU PONT DE NEMOURS & CO E I.
XX
XX Allen SM, Hitz WD, Kinney AJ, Tingey SV;
XX
XX WPI; 2002-453364/48.
DR P-PSDB; AAU97207.
XX
XX
PT New nucleic acid encoding plant sugar-transport proteins, useful for
PT preparing transgenic plants with altered carbohydrate distribution.
XX
XX
PS Claim 3; Col 47-50; 54pp; English.
XX
XX
CC The present invention relates to the isolation of plant polynucleotide
CC sequences encoding an Arabidopsis thaliana-like sugar transport protein
CC or Beta vulgaris-like sugar transport protein. The polynucleotide
CC sequences are useful for altering the level of sugar transport proteins
CC in plants, i.e. for control of carbohydrate transport and distribution in
CC plant cells, e.g. during grain filling of annual field crops (e.g. corn,
CC rice, soybeans, and wheat), and, for studying carbohydrate flows and
CC sugar transport. The polynucleotide sequences can also be used to isolate
CC cDNA sequences and genes that encode homologues of the new proteins. The
CC present sequence representing a contig assembled from various wheat cDNA
CC clones encodes a portion of an Arabidopsis thaliana-like sugar transport
CC protein
XX
SQ Sequence 1487 BP; 355 A; 320 C; 373 G; 439 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 7.77e-115 Length: 1487
Score: 1424.00 Matches: 281
Percent Similarity: 89.97% Conservative: 24
Best Local Similarity: 82.89% Mismatches: 32
Query Match: 34.94% Indels: 2
DB: 6 Gaps: 2

US-10-051-909-32 (1-800) x ABK51968 (1-1487)

QY 462 GluGlyGlyAspGlyValSerSerThrAspIleGlyGlyTrpGlnLeuAlaTrpLys 481
Db 12 GAGGGTGGGAGGCGAGTCAGCAGCAGCTGGTATTGGTGGGGGTGGCAACTCGCATGGAAA 71

QY 482 TrpSerGluLysGluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyrLeu 501
Db 72 TGGTCGGAGCGACAGGCGAGGCTGGCAAGAGGAGGAGGCTTCAAAGAATCTACTTG 131

QY 502 HisGlnGluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGlyGly 521
Db 132 CACCAAGAGGGGGTGGCGGCACTCAAGAGGGGCTGTGTTTCACTTCCTGGTGGGGT 191

QY 522 AspValPheGlu---GlySerGluPheValHisAlaAlaLeuValSerGlnSerAla 540
Db 192 GATGCCACGCAAGGGGCGAGTGGGTTTATACATGCTGCTGCTTTGGTAAGCCACTCGGCT 251

QY 541 LeuPheSerLysGlyLeuAlaGluProArgMet---SerAspAlaAlaMetValHisPro 559
Db 252 CTTTACTCCAAGGATCTTATGAAGAGCGGTATGGCGCGGTCCAGCCATGATTCATCCA 311

QY 560 SerGluValAlaAlaLysGlySerArgTrpLysAspLeuPheGluProGlyValArgArg 579
Db 312 TTGGAGGCGAGCTCCCAAGGTTCAATCTGGAAGATCTGTTGAACCTGGTGTGAGGCGT 371

QY 580 AlaLeuLeuValGlyValGlyIleGlnIleLeuGlnPheAlaGlyIleAsnGlyVal 599
Db 372 GCATTGTCGTCGGTGTGGAATTCAGATGCTTCAGCAGTTTCTGGAATAAATGGAGTT 431

QY 600 LeuTyrThrProGlnIleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPhe 619
Db 432 CTCTACTATACCTCTCAAAATCTGGAGCAAGCTGGTGGCTGTTCTTCTTCCAATCTT 491

QY 620 GlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeu 639
Db 492 GGCTCAGTTCAGCATCAGCATCCATCTTGATCAGTTCTCTCACCACCTTACTCATGCTC 551

QY 640 ProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeu 659
Db 552 CCAAGCATGGTGTAGCCATGAGACTTATGGATATATCTGGAAGAGGTTTCTGCTACTG 611

QY 660 GlyThrIleProIleLeuIleAlaSerLeuValIleLeuValValSerAsnLeuIleAsp 679
Db 612 GGCACAATTCCTCATCTGATAGCATCCCTAATTGTTGGGTGGTGGTCAATGTATCAAC 671

QY 680 LeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyrPheCysCys 699
Db 672 TTGAGTACGGTGGCCCGCTGTGCTCTCCACAGTAGCGTCAATTGCTACTTCTGCTGC 731

QY 700 PheValMetGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePheProThrArg 719
Db 732 TTTGTCATGGGCTTTGGCCCGATCCCAACATTCATGTGCAGAGATTTTCCCCACCAGA 791

QY 720 ValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIleGlyAspIleIleVal 739
Db 792 GTCCGTGGTGTCTGCATCGCTATTTCGCCCTCACATTCTGGATTGTGACATATTGTT 851

QY 740 ThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyr 759
Db 852 ACCTACAGCCTGCCTGTGATGCTGAATGCTATTGCTAGCGGTGCTTTGGTATATAT 911

QY 760 AlaValValCysLeuIleSerPheValPheValPheLeuLysValProGluThrLysGly 779
Db 912 GCAGTCGTTTGTGTCATTCCTTTGTGTTCTGCTACCTAAAGGTCCAGAGACAAAGGGC 971

QY 780 MetProLeuGluValIleThrGluPhePheAlaValGlyAlaLysGlnAlaAla 798
Db 972 ATGCCCTCGAGGTTCATCACCAGATTCTTTGCGGTGGGGCGAAGCAAGCGCAGGCC 1028

Search completed: October 13, 2004, 12:23:49
Job time : 1070.57 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 13, 2004, 11:54:50 ; Search time 169.106 Seconds
(without alignments)
3362.578 Million cell updates/sec

Title: US-10-051-909-32
Perfect score: 4075
Sequence: 1 IRSGSWLAVQTPTPLDRR.....PLEVITEFFAVGAKQAAKA 800

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 824507 seqs, 355394441 residues
Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USFTO_spool_h/US10051909/runat_13102004_123337_19931/app_query.fasta_1.1678
-DB=Issued Patents_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=Bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10051909 @CGN 1 116 @runat_13102004_123337_19931 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents_NA:
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	3517	86.3	2824	3	US-09-291-922-1 Sequence 1, Appli
2	2674	65.6	2601	3	US-09-291-922-7 Sequence 7, Appli
3	1670.5	41.0	1692	3	US-09-291-922-9 Sequence 9, Appli
4	1424	34.9	1487	3	US-09-291-922-13 Sequence 13, Appl
5	943	23.1	1009	3	US-09-291-922-15 Sequence 15, Appl
6	562	13.8	870	3	US-09-291-922-5 Sequence 5, Appli
7	496	12.2	1853	3	US-09-291-922-23 Sequence 23, Appl
8	494.5	12.1	2017	3	US-09-291-922-21 Sequence 21, Appl
9	491.5	12.1	1914	3	US-09-291-922-19 Sequence 19, Appl
10	489	12.0	1872	3	US-09-291-922-27 Sequence 27, Appl
11	463.5	11.4	2089	3	US-09-291-922-25 Sequence 25, Appl
12	447	11.0	1431	4	US-09-489-039A-4762 Sequence 4762, Ap

ALIGNMENTS

RESULT 1

US-09-291-922-1
; Sequence 1, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 2824
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (29)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (622)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (636)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (638)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (669)

13	440.5	10.8	1545	4	US-09-489-039A-4731	Sequence 4731, Ap
14	427	10.5	1506	4	US-09-489-039A-4560	Sequence 4560, Ap
15	426	10.5	1566	4	US-09-489-039A-2378	Sequence 2378, Ap
16	418.5	10.3	1668	4	US-09-614-221A-420	Sequence 420, App
17	412	10.1	4411529	3	US-09-103-840A-1	Sequence 1, Appli
18	368	9.0	4403765	3	US-09-103-840A-2	Sequence 2, Appli
19	355	8.7	2856	4	US-09-643-597-135	Sequence 135, App
20	355	8.7	2856	4	US-09-480-884A-135	Sequence 135, App
21	355	8.7	2856	4	US-09-542-615A-135	Sequence 135, App
22	355	8.7	2856	4	US-09-606-421B-135	Sequence 135, App
23	355	8.7	2856	4	US-09-221-107-135	Sequence 135, App
24	355	8.7	2856	4	US-09-466-396A-135	Sequence 135, App
25	355	8.7	2856	4	US-09-476-496A-135	Sequence 135, App
26	355	8.7	2856	4	US-09-630-940B-135	Sequence 135, App
27	354.5	8.7	3000	2	US-08-928-692-9	Sequence 9, Appli
28	354.5	8.7	3000	3	US-09-339-972-9	Sequence 9, Appli
29	352.5	8.7	1626	4	US-09-614-221A-521	Sequence 521, App
30	346	8.5	5228	4	US-09-919-039-216	Sequence 216, App
31	343	8.4	1776	4	US-09-679-686B-11	Sequence 11, Appl
32	338.5	8.3	1943	4	US-09-774-528-168	Sequence 168, App
33	334.5	8.2	5227	4	US-09-919-172-79	Sequence 79, Appl
34	332.5	8.2	2592	3	US-09-591-025-8	Sequence 8, Appli
35	328.5	8.1	2592	4	US-09-894-927B-8	Sequence 8, Appli
36	322	7.9	1659	4	US-09-248-796A-6610	Sequence 6610, Ap
37	313.5	7.7	1704	4	US-09-614-221A-534	Sequence 534, App
38	312	7.7	1675	4	US-09-679-686B-1	Sequence 1, Appli
39	312	7.7	1695	4	US-09-614-221A-443	Sequence 443, App
40	307.5	7.5	1704	4	US-09-614-221A-96	Sequence 96, Appl
41	302	7.4	2298	4	US-09-248-796A-5965	Sequence 5965, Ap
42	301.5	7.4	3915	4	US-09-023-655-1104	Sequence 1104, Ap
43	293	7.2	987	4	US-09-710-279-1941	Sequence 1941, Ap
44	293	7.2	3065	4	US-09-710-279-3697	Sequence 3697, Ap
45	291.5	7.2	1752	4	US-09-679-686B-17	Sequence 17, Appl

FEATURE:
NAME/KEY: unsure
LOCATION: (771)
FEATURE:
NAME/KEY: unsure
LOCATION: (822)
FEATURE:
NAME/KEY: unsure
LOCATION: (856)
FEATURE:
NAME/KEY: unsure
LOCATION: (889)
FEATURE:
NAME/KEY: unsure
LOCATION: (944)
US-09-291-922-1

Alignment Scores:

Pred. No.: 0
Score: 3517.00
Percent Similarity: 95.18%
Best Local Similarity: 91.57%
Query Match: 86.31%
DB: 3

US-10-051-909-32 (1-800) x US-09-291-922-1 (1-2824)

Table with 3 columns: QY/Db, sequence, and position. It lists amino acid sequences for various proteins, including MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTyr, and others, with their corresponding positions from 56 to 838.

Table with 3 columns: QY/Db, sequence, and position. It continues the list of amino acid sequences from position 274 to 1918, including SerLeuLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIle, and others.

QY 634 ThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGly 653
 Db 1978 ACTACTTACTAATGCTTCCTTGCATTTGGCTTTGCCATGCTGCTTATGGATCTTTCCGGA 2037
 QY 654 ArgArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuVal 673
 Db 2038 AGAAGGTTTTTGTGCTAGGCACAATTCCAATCTTGATAGCATCTCTAGTATCTCTGGTT 2097
 QY 674 ValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerVal 693
 Db 2098 GTGTCCAATCTAATTTGATTTGGGTACACTAGCCCATGCTTTGTCTCTCCACCATCAGTGT 2157
 QY 694 IleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAla 713
 Db 2158 ATCGTCTACTTCTGCTGCTTCGTTATGGGATTTGGTCCCATCCCAACATTTTATGTGCA 2217
 QY 714 GluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrp 733
 Db 2218 GAGATCTTTCCAAACCAGGTTGCTGGCTCTGTATGCCATTTGTGCCTTTACATTTCTGG 2277
 QY 734 IleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAla 753
 Db 2278 ATCGGAGATATCATCGTCACTACAGCCTTCCCTGTGATGCTGAATGCTATTGGACTGGCG 2337
 QY 754 GlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeuLys 773
 Db 2338 GGTTGTTTCAGCATATATGAGTGCATGATGCTTGATTTCTTTGTGTTCTCTTCTTAAAG 2397
 QY 774 ValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAla 793
 Db 2398 GTCCCTGAGACAAAGGGGATGCCCTTGAGGTTATTACCGAATTCITTTGCAGTTTGGTGG 2457
 QY 794 LysGlnAlaAlaAlaLysAla 800
 Db 2458 AAGCAAGCGGCTGCAAAAGCC 2478

RESULT 2

US-09-291-922-7
 ; Sequence 7, Application US/09291922
 ; Patent No. 6383776
 ; GENERAL INFORMATION:
 ; APPLICANT: Allen, Steve
 ; APPLICANT: Hitz, Bill
 ; APPLICANT: Kinney, Tony
 ; APPLICANT: Tingey, Scott
 ; TITLE OF INVENTION: Plant Sugar Transport Proteins
 ; FILE REFERENCE: BB-1163
 ; CURRENT APPLICATION NUMBER: US/09/291,922
 ; CURRENT FILING DATE: 1999-04-14
 ; EARLIER APPLICATION NUMBER: 60/083,044
 ; EARLIER FILING DATE: April 24, 1998
 ; NUMBER OF SEQ ID NOS: 30
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 7
 ; LENGTH: 2601
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 US-09-291-922-7

Alignment Scores:

Pred. No.: 1.7e-272 Length: 2601
 Score: 2674.00 Matches: 522
 Percent Similarity: 82.20% Conservative: 92
 Best Local Similarity: 69.88% Mismatches: 115
 Query Match: 65.62% Indels: 18
 DB: 3 Gaps: 8

US-10-051-909-32 (1-800) x US-09-291-922-7 (1-2601)

QY 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrp 75
 Db 175 ATGAAAGGTGCCGTCCTTGTGCTATTGCCGCTTCCATTGGTAATTCCTCCAAAGGATGG 234

QY 76 AspAsnAlaThrIleAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsn 95
 Db 235 GATAATGCTACCATCGCCGGGCTAATGGTTACATTAAGAAAGACCTTGTCTTTGGGAACA 294
 QY 96 GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115
 Db 295 -----ACTATGAAAGGCTTGTGTGGGCATGTCCCTGATGGAGCAACGGTAATCACC 348
 QY 116 ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer 135
 Db 349 ACATGCTCTGTCCTATAGCGGATTTGGCTCGGTGGCGACCATGATGATAATCTCATCT 408
 QY 136 IleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyrValLeuLeu 155
 Db 409 GTGCTCTATTTCTTGGTGGTTTGGTGATGCTGTGGTCCCAAAATGTGTATGTGTGTGC 468
 QY 156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr 175
 Db 469 TTGGCGAGGCTACTTGATGGATTTGGGATTGGCCTTGTGTGACTCTTGTCCCGGCTAT 528
 QY 176 IleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSer 195
 Db 529 ATATCTGAACCGGCGCTCTGAATAAGGGGTCGTTGAATACGCTTCTCCTCAGTTTCA 588
 QY 196 GlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSer 215
 Db 589 GGCTCTGGAGGAATGTTTGTGCTACTGTATGGTTTGGCATGTCTATTGATGCCGCG 648
 QY 216 ProAspTyrArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeu 235
 Db 649 CCTAGCTGGAGGCTCATGCTTGGGTTCTGTCTATTCTCTCTCTTGTATTTTGCATTG 708
 QY 236 ThrIlePheTyrLeuProGluSerProArgTyrLeuValSerLysGlyArgMetAlaGlu 255
 Db 709 ACCATTTTTTCTTGGCCGAGTCTCCTCGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 768
 QY 256 AlaLysLysValLeuGlnLysLeuArgGlyLysAspAspValSerGlyGluLeuSerLeu 275
 Db 769 GCTAAGAAGTCTCCAAAGATTGCGGGAAGGAGGAGGATGTGTGAGGAGATGGCATTG 828
 QY 276 LeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIleGlyPro 295
 Db 829 CTGGTTGAAGTCTCGGGATTGGGGTGATACATCTATCGAAGAGTACATAATTGSCCT 888
 QY 296 AlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyr 315
 Db 889 GCTGACGATGTGGCTGATGGTTCATGAACATGCAACAGAGAAAGATAAATTCGATTAT 948
 QY 316 GlyProGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGly 335
 Db 949 GGATCCCAAGCAGGCTTCTTCTGGTTATCAAAACCTGTCACAGAGTTCTATTGGC 1008
 QY 336 SerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValProLeuMet 355
 Db 1009 -----CTTGGCTCACACCATGGAAGCATCATCAACCAAGCATGCCCTCATG 1056
 QY 356 AspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly----- 373
 Db 1057 GATCCTCTGTGACACTGTTTGTAGCATTCATGAGAGCTCCCGAGACAGGAGCAAGA 1116
 QY 374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393
 Db 1117 GGAAGCATCGAAGCACTCTGTTTCCAAATTTTGGAGCATGTTTCAGCACTGTGAGCCG 1176
 QY 394 HisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAla 413
 Db 1177 CATGCTAAAATGAACAATGGGATGAAGAAAGCTTTACAAAGGAACGTTGAGGACTACATG 1236
 QY 414 SerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGln 433
 Db 1237 TCAGATGCAACCCCGTGGGACTCCGATGATAATTTGACAGTCTTTTAACTCTCAGCCAA 1296
 QY 434 AlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeu--- 452

Db 1297 ACAACAAGCCCTTGAA---AAAGACTTACCTCCTCCTCCTCCATGGCAGTATCCTTGGC 1353
QY SerMetArgArgGlnThrLeuLeuGlyGluGly---GlyAspGlyValSerSerThrAsp 471
Db 1354 AGCATGAGGCGTCACAGTAGTCTCATGCAAGGGTCAGGTGAGCAAGGTGGTAGTACAGT 1413
QY IleGlyGlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArg 491
Db 1414 ATTGGTGGTGGCTGGCAACTGGCATGGAATGGACTGATAA---GGTGAGGATGGAAAA 1470
QY LysGluGlyGlyPheLysArgValTrpLeuHisGlnGluGlyValProGlySerArgArg 511
Db 1471 CAACAAGGAGGGTTAAAGGATTATTATACATGAGGAGGAGTTTCTGCATCTCGTCGT 1530
QY GlySerIleValSerLeuProGlyGlyGlyGlyAspValPheGluGlySerGluPheValHis 531
Db 1531 GGATCCATTGTATCGATTCCCGGTGAAGGC-----GAATTGTCCAG 1572
QY AlaAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMet 551
Db 1573 GCTGCTGCCTTGGTAAGCCCAACCGCTCTTACTCCAAGGAGCTTATTGATGGACACCCA 1632
QY SerAspAlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAsp 571
Db 1633 GTTGGCGCTGCAATGGTTACCCATCTGAGACAGCTTCAAGGGGCCAAGTTGGAAGCT 1692
QY LeuPheGluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGln 591
Db 1693 CTTCTTGAACCAAGGGTTAAGCATGCTATTGGTTGGAGTTGGAATACAATACTTCAG 1752
QY GlnPheAlaGlyIleAsnGlyValLeuTyrTrpThrProGlnIleLeuGluGlnAlaGly 611
Db 1753 CAGTTTTCAGGGATAAATGGGTTCTATATTACACACTCAATCCTTTGAAGAGCGCGT 1812
QY ValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSer 631
Db 1813 GTTGAAGTTCTTCTTTTCAGATATAGGCATTGGCTCAGAGTCGGCATCATCTTATCAGT 1872
QY SerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeu 651
Db 1873 GCTTTCACAACCTTCTTGATGTTCTCCTGTATAGCGTAGCCATGAAGTCATGGATGTT 1932
QY SerGlyArgArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIle 671
Db 1933 TCAGGCAGAAGGCAGTTGCTACTTACTACAAATCCCGTCTGATTGTGTCACTCATTATT 1992
QY LeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrVal 691
Db 1993 TTGGTCATTGGAAGCCTGGTAAATTTTGGCAATGTGCGCCCATGCAGCAATCTCAACAGTA 2052
QY SerValIleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeu 711
Db 2053 TGCCTTGTGGTTTATTCTGTCTGCTTGTGATGGGTTATGGACCAATTCACCATCCTT 2112
QY CysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThr 731
Db 2113 TGCTCAGAGATTTTCCCACTAGGTGCGTGGCTCTGCTATGCTATCTGTGCTATTAGTG 2172
QY PheTrpIleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGly 751
Db 2173 TTCTGGATTGGAGACATCATCATACATACTGCTGCTGCTGTGATGCTCGGCTCTTTAGGA 2232
QY LeuAlaGlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPhe 771
Db 2233 CTTGGTGGTGTATCGCCATTACGCAGTTGTTTGTTCATCTCGTGGATATTGTGTTT 2292
QY LeuLysValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaVal 791
Db 2293 TTGAAGGTTCCAGAAACAAAGGCGCATGCCCTTGAAGTCATCTCTGAATCTTTTCTGTT 2352
QY GlyAlaLysGlnAlaAla 798

Db 2353 GGAGCAAAAGCAGCGTCTTCT 2373
RESULT 3
US-09-291-922-9
; Sequence 9, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 9
; LENGTH: 1692
; TYPE: DNA
; ORGANISM: Glycine max
US-09-291-922-9
Alignment Scores:
Pred. No.: 1.56e-166 Length: 1692
Score: 1670.50 Matches: 333
Percent Similarity: 79.75% Conservative: 57
Best local Similarity: 68.10% Mismatches: 84
Query Match: 40.99% Indels: 15
DB: 3 Gaps: 8
US-10-051-909-32 (1-800) x US-09-291-922-9 (1-1692)
QY 308 AspLysGluGlnIleThrLeuTyrGlyProGluGluGlyGlnSerTrpIleAlaArgPro 327
Db 21 GAAAAAGATCAAAATTAAGTTGTATGGACCAGAACAGGCCAGTCTCTGGTTCCTAGACCT 80
QY 328 SerLysGlyProIleMetLeuGlySerValLeuSerLeuAlaSerArgHisGlySerMet 347
Db 81 GTTGCTGGACCAAAATTCCTGTGGC-----CTTGATCTAGGAAAGGAGCATG 128
QY 348 ValAsnGlnSerValProLeuMetAspProIleValThrLeuPheGlySerValHisGlu 367
Db 129 GCAAAATCCAAGC---AGTCTAGTGGACCCCTCTAGTGACCCCTCTTTGGTAGTGTACATGAG 185
QY 368 AsnMetProGlnAlaGlyGlySerMetArgSerThrLeuPheProAsnPheGlySerMet 387
Db 186 AAGCTCCCAGAAACAGGA-----AGCACCCCTTTTCCACACTTTGGGAGTATG 233
QY 388 PheSerValThrAspGlnHisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArg 407
Db 234 TTCAGTGTGGGGAAATCAGCCCAAGGATGAAGATTGGGATGAGGAAAGCCTAGCCAGA 293
QY 408 AspAspGluGluTyrAlaSerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSer 427
Db 294 GAGGGTGTATGATTATGTCTCTGAT-----GCTGGTGTATCTGTATGACAAATTTGCAGAGT 347
QY 428 ProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHis 447
Db 348 CCATTGATCTCAGCTCAACCAACGAGTCTGGAT---AAGGACATACCTCTCATGCCCCAT 404
QY 448 ArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeu---GlyGluGlyGlyAspGly 466
Db 405 AGTAACCTTGCA---AGCATGAGGCAAGGTAGTCTTTTACATGCAAAATTCAGGAGAACCC 461
QY 467 ValSerSerThrAspIleGlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGlu 486
Db 462 ACTGGTAGTACTGGGATTGGTGGTGGTGGCAGCTAGCATGCAAAATGGTCTGAAAGAGAG 521
QY 487 GlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGlyVal 506

Db 522 GGGCCAGATGGAAGAAGAGGTGGCTTCAAGAGAAATATATTACACCAAGATGGTGT 581
QY 507 ProGlySerArgGlySerIleValSerLeuProGlyGlyGlyAspValPheGluGly 526
Db 582 TCTGGATCTAGACGTGGTCTGTGGTTTCACTCCCT--GGCGGTGATTACCAACTGAC 638
QY 527 SerGluPheValHisAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeu 546
Db 639 AGTGAGGTGTACAGGCTGCTGCTGTGGTGAAGTCAGCCTGCCCTTTATATAGGACCTT 698
QY 547 AlaGluProArgMetSerAspAlaAlaMetValHisProSerGluValAlaAlaLysGly 566
Db 699 ATGGGTCAACGGCCAGTTGGACCAGCTATGATTCTCCCTCTGAACAATTTGCAAAAGGG 758
QY 567 SerArgTrpLysAspLeuPheGluProGlyValArgAlaLeuValGlyValGly 586
Db 759 CCAAGTTGGAGTGATCTTTTGAACCTGGGTGAAGCATGCATGATTGTGGGGTGGGA 818
QY 587 IleGlnIleLeuGlnPheAlaGlyIleAsnGlyValLeuTyrThrProGlnIle 606
Db 819 ATGCAAAATCTTCAGCAGTTCTCTGGTATAAATGGGTCCCTACTATACGCCTCAAAAT 878
QY 607 LeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAla 626
Db 879 CTTGAGCAGGAGGTGTGGTTATCTTCTTTCAAGCCTAGGCTTGGTTCTACTTCTTCA 938
QY 627 SerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMet 646
Db 939 TCCTTTCTTATTAGTGGGTGACAACTTGTGATGCTTCTTGTATAGCCATTGGCCATG 998
QY 647 LeuLeuMetAspLeuSerGlyArgPheLeuLeuLeuGlyThrIleProIleLeuIle 666
Db 999 AGGCTCATGGATATTTCAGGCAGAGGACTTTGTCTGCTCAGTACATCCCGTCTCTAATA 1058
QY 667 AlaSerLeuValIleLeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAla 686
Db 1059 GCAGCTCTTCTCATATTAGTCTCTGGGAAGTCTGTGGATTGGGATCCACTGCAAAATGCA 1118
QY 687 LeuLeuSerThrValSerValIleValTyrPheCysCysPheValMetGlyPheGlyPro 706
Db 1119 TCAATCTCAACCATTAGTGTATTGTCTATTCTTCTTGTTCATGGGATTGGACCA 1178
QY 707 IleProAsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIleAla 726
Db 1179 ATTCTTAATATATTGTGCAGAGATCTTCCCACTCGAGTTCGTGGTCTCTGCATTGCT 1238
QY 727 IleCysAlaPheThrPheTrpIleGlyAspIleIleValThrTyrSerLeuProValMet 746
Db 1239 ATTTGTGCCCTTACCTTTTGGATCTGTGATATCATTTGTACCTACACACTCCCAGTTATG 1298
QY 747 LeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeuIleSer 766
Db 1299 CTCAATCTGTAGGCCTCGCTGGTGTTTTGGTATTTATGCTGTCGTGCTTCATAGCA 1358
QY 767 PheValPheValPheLeuLysValProGluThrLysGlyMetProLeuGluValIleThr 786
Db 1359 TGGGTGTTTGTCTTTTGAAGTTCCAGAAACCAAGGGCATGCCACTGGAAAGTGATCATT 1418
QY 787 GluPhePheAlaValAlaLysGln 795
Db 1419 GAGTTCTTCTGTGCGGAGCAAAACAG 1445

RESULT 4
US-09-291-922-13
; Sequence 13, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163

; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 13
; LENGTH: 1487
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-09-291-922-13
Alignment Scores: 1.59e-140 Length: 1487
Pred. No.: 1424.00 Matches: 281
Score: 89.97% Conservative: 24
Percent Similarity: 82.89% Mismatches: 32
Best Local Similarity: 34.94% Indels: 2
Query Match: 3 Gaps: 2
DB:
US-10-051-909-32 (1-800) x US-09-291-922-13 (1-1487)
QY 462 GluGlyGlyAspGlyValSerSerThrAspIleGlyGlyGlyTrpGlnLeuAlaTrpLys 481
Db 12 GAGGGTGGGAGGCGAGTCAGCAGCACTGGTATTGGTGGGGGTGGCAACTCGCATGGAAA 71
QY 482 TrpSerGluLysGluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyrLeu 501
Db 72 TGGTCGGAGCGACAAAGCGGAGGATGGCAAGAGGAAGAGGCTTCAAAGAATCTACTTG 131
QY 502 HisGlnGluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGlyGly 521
Db 132 CACCAAGAGGGGTGGCCGACTCAAGAAAGGGGCTCTGTTGTTTCACTTCTCTGGTGGGGT 191
QY 522 AspValPheGlu---GlySerGluPheValHisAlaAlaAlaLeuValSerGlnSerAla 540
Db 192 GATGCCACGCAAGGGGCGAGTGGTTTATACATGCTGCTGCTTTGGTAAGCCACTCGGCT 251
QY 541 LeuPheSerLysGlyLeuAlaGluProArgMet---SerAspAlaAlaMetValHisPro 559
Db 252 CTTTACTCCAAGGATCTTATGGAAGAGCGGTATGGCGCGCGGTCCAGCCATGATCATCCA 311
QY 560 SerGluValAlaAlaLysGlySerArgTrpLysAspLeuPheGluProGlyValArgArg 579
Db 312 TTGGAGGCAGCTCCCAAGGTTCAATCTGGAAGATCTGTTTGAACCTGGTGTGAGGCGT 371
QY 580 AlaLeuLeuValGlyValGlyIleGlnIleLeuGlnPheAlaGlyIleAsnGlyVal 599
Db 372 GCATTGTTCTGTCGGTGTGGAAATTCAGATGCTTCAGCAGTTCAGTGGTTCCTTCTTCCAATCTT 431
QY 600 LeuTyrTyrThrProGlnIleLeuGluAlaGlyValAlaValIleLeuSerLysPhe 619
Db 432 CTCTACTATCTCCTCAAAATCTGGAGCAAGCTGGTGGTGGTGTCTTCTTCTTCCAATCTT 491
QY 620 GlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeu 639
Db 492 GGCTCAGTTCAGCATCAGCATCCATCTTGATCAGTTCCTCACCACCTTACTCATGCTC 551
QY 640 ProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeu 659
Db 552 CCAAGCATTTGGTAGCCATGAGACTTATGGATATATCTGGAAGAAGGTTTCTGCTACTG 611
QY 660 GlyThrIleProIleLeuIleAlaSerLeuValIleLeuValValSerAsnLeuIleAsp 679
Db 612 GGCACAAATTCCTTCTGATAGCATCCCTAAATTTGGGTGGTGGTCAATGTTATCAAC 671
QY 680 LeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyrPheCysCys 699
Db 672 TTGAGTACGGTGGCCCAACCGCTGTGCTCTCCACAGTTAGGCTCATTTCTTCTTCTGCTGC 731
QY 700 PheValMetGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePheProThrArg 719
Db 732 TTTGTATGGGCTTTTGGCCCGATCCCAACATTTCTATGTGAGAGATTTTCCCCCACCAGA 791

QY 720 ValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIleGlyAspIleVal 739
 Db 792 GTCCGTTGGTCTGCATCGCTATTGCGCCCTCACATTCTGGATTGTGACATTATTGTT 851
 QY 740 ThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyr 759
 Db 852 ACCTACAGCTGCTGTGATGCTGAATGCTATTGGTCTAGCGGTGTCTTTGGTATATAT 911
 QY 760 AlaValValCysLeuIleSerPheValPheValPheLeuLysValProGluThrLysGly 779
 Db 912 GCAGTCTGTTGCTGCTATGCTTGTGCTTGTGCTACCTAAAGGTCCCAGAGACAAAGGC 971
 QY 780 MetProLeuGluValIleThrGluPhePheAlaValGlyAlaLysGlnAlaAla 798
 Db 972 ATGCCCCCTCGAGGTCTATCACCAGTCTTTTGGGTTGGGCGAAGCAAGCGCAGGCC 1028
 RESULT 5
 US-09-291-922-15
 ; Sequence 15, Application US/09291922
 ; Patent No. 6383776
 ; GENERAL INFORMATION:
 ; APPLICANT: Allen, Steve
 ; APPLICANT: Hitz, Bill
 ; APPLICANT: Kinney, Tony
 ; APPLICANT: Tingey, Scott
 ; TITLE OF INVENTION: Plant Sugar Transport Proteins
 ; FILE REFERENCE: BB-1163
 ; CURRENT APPLICATION NUMBER: US/09/291,922
 ; CURRENT FILING DATE: 1999-04-14
 ; EARLIER APPLICATION NUMBER: 60/083,044
 ; EARLIER FILING DATE: April 24, 1998
 ; NUMBER OF SEQ ID NOS: 30
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 15
 ; LENGTH: 1009
 ; TYPE: DNA
 ; ORGANISM: Triticum aestivum
 US-09-291-922-15
 Alignment Scores:
 Pred. No.: 7.15e-90 Length: 1009
 Score: 943.00 Matches: 179
 Percent Similarity: 91.89% Conservative: 25
 Best Local Similarity: 80.63% Mismatches: 18
 Query Match: 23.14% Indels: 0
 DB: 3 Gaps: 0
 US-10-051-909-32 (1-800) x US-09-291-922-15 (1-1009)
 QY 574 GluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnPhe 593
 Db 2 GAACCTGGAGTGAAGCATGCACTGTTGTTGGCATAGGATTACAGATCCTGAGCAGTTT 61
 QY 594 AlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyValAla 613
 Db 62 GCGGGTATCAATGGAGTCTCTACTACACACCTCAGATACCTTGGCAAGCAGGTGTCGGG 121
 QY 614 ValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeu 633
 Db 122 GTTCTCTATCAACATTGGACTAAGCTCTTCTCAGCATCTATTCTTTATTAGTGCCTTG 181
 QY 634 ThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGly 653
 Db 182 ACAACCTTGCTGATGCTTCCAGCATTTGGCATCGCCATGAGACTCATGGATATGTCAGGA 241
 QY 654 ArgArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuVal 673
 Db 242 AGAAGGTTTCTTCTCTTCAACAATCCCTGTCTGTAGTAGCGGTAGTGTCTTGGTT 301
 QY 674 ValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerVal 693
 Db 302 TTAGTGAATGTTCTGGATGTCGGAACCATGGTGCAGCGCTGCGCTCTCTCAGCATCAGCGTC 361

QY 694 IleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAla 713
 Db 362 ATCGTCTATTCTGCTTCTTCGTCATGGGGTTTGGGCCATATCCCAATATTCTCTGCGCG 421
 QY 714 GluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrp 733
 Db 422 GAGATTTTCCCACTCTCTGTCGTCGTCATCTGCATAGCCATCTGCGCGCTAACCTTCTGG 481
 QY 734 IleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAla 753
 Db 482 ATCGGCGACATCATCGTGACATACACTCTCCCCCGTGTATGCTCAATGCCATTGGTCTCGCT 541
 QY 754 GlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeuLys 773
 Db 542 GGAGTCTTCGGCATATATGCCATCGTTTGTGTACTAGCCCTTGTATTCTCTACATGAAG 601
 QY 774 ValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAla 793
 Db 602 GTCCCTGAGACAAAGGGCATGCCCTGGAGGTCTATCACCAGTCTTCTCTCTGTCGGGCA 661
 QY 794 LysGln 795
 Db 662 AAGCAG 667
 RESULT 6
 US-09-291-922-5
 ; Sequence 5, Application US/09291922
 ; Patent No. 6383776
 ; GENERAL INFORMATION:
 ; APPLICANT: Allen, Steve
 ; APPLICANT: Hitz, Bill
 ; APPLICANT: Kinney, Tony
 ; APPLICANT: Tingey, Scott
 ; TITLE OF INVENTION: Plant Sugar Transport Proteins
 ; FILE REFERENCE: BB-1163
 ; CURRENT APPLICATION NUMBER: US/09/291,922
 ; CURRENT FILING DATE: 1999-04-14
 ; EARLIER APPLICATION NUMBER: 60/083,044
 ; EARLIER FILING DATE: April 24, 1998
 ; NUMBER OF SEQ ID NOS: 30
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 5
 ; LENGTH: 870
 ; TYPE: DNA
 ; ORGANISM: Oryza sativa
 US-09-291-922-5
 Alignment Scores:
 Pred. No.: 1.25e-49 Length: 870
 Score: 562.00 Matches: 103
 Percent Similarity: 93.75% Conservative: 17
 Best Local Similarity: 80.47% Mismatches: 8
 Query Match: 13.79% Indels: 0
 DB: 3 Gaps: 0
 US-10-051-909-32 (1-800) x US-09-291-922-5 (1-870)
 QY 669 LeuValIleLeuValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeu 688
 Db 8 GTTCTAACCTTGATTCTGCTCAATATTCTGGATGTTGGGACCATGGTTTCATGCTCACTG 67
 QY 689 SerThrValSerValIleValTyrPheCysCysPheValMetGlyPheGlyProIlePro 708
 Db 68 TCCACAGTCAGTGTCTACTTCTTCTGCTTCTTGTCTATGGGTTTCGGGCTTATTTCCA 127
 QY 709 AsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCys 728
 Db 128 AACATTCTCTGTGCAGAGATTTTCCCGACCATGTTGTTGTCATCTGTCATAGCCATCTGT 187
 QY 729 AlaPheThrPheTrpIleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsn 748
 Db 188 GCGCTAACATTCTGGATCGGTGATATCATTTGTGACATACACCTTCCCCCGTGTGATGCTCAAC 247

QY 556 MetValHisProSerGluValAlaAlaLysGlySerArgTyrLysAspLeuPhe----- 573
Db 986 -----TGAAGAGAGCTCTTCTCTAT 1006
QY 574 ---GluProGlyValArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGln 592
Db 1007 CCAACGCCGCAATTCTGTCACATCGTAATCGCTGCCCTTGGTATTCACTTCTTCCAACAA 1066
QY 593 PheAlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyVal 612
Db 1067 GCGTCGGCGGTAGACGCCGCTGTTTGTACAGCCCGCAGGATCTTCGAAAGGCTGGGATT 1126
QY 613 AlaValIleLeuSerLysPheGlyLeuSerSerAlaSerIleLeuIleSerSer 632
Db 1127 ACAAACGACACGATAGCTT-----CTTGCAACCGTGGCGCTGGATT 1171
QY 633 LeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSer 652
Db 1172 GTTAAGACCGTGTTCATCTTGGCG-----GCTACGTTTACGTTGGACCGCGTG 1219
QY 653 GlyArgArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeu 672
Db 1220 GGTGCTGCTCCGTTGTTATTGTTCTAGTGTGGCGGCATGGTCTCTCGCTTCTCACGCTT 1279
QY 673 ValValSer---AsnLeuIleAspLeuGly-----ThrLeuAlaHisAlaLeuLeuSer 689
Db 1280 GCGATCAGCCTCACTGTTATTGATCATTCGGAGAGGAAATTAATGTGGCGCTGGATCG 1339
QY 690 ThrValSer---ValIleValTyrPheCysCysPheValMetGlyPheGlyProIlePro 708
Db 1340 AGCATAGCCATGGTGTGGCTTACGTGGCCACGTTCTCCATCGGTGCGGTCGCCATCAG 1399
QY 709 AsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCys 728
Db 1400 TGGGTCTATAGTTCTGAGATCTTCCCGTTGAGGCTGCGGCGCARGGTGCGGCGCGGGA 1459
QY 729 AlaPheThrPheTyrIleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsn 748
Db 1460 GTTGGCGTGAATAGGACCCTAGCGCGGTGTCTCAATGACTTTTCTGTCCCTCACTAGA 1519
QY 749 AlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeuIleSerPheVal 768
Db 1520 GCCATCACTATTGGTGAGCTTCTTCTTATTGTCATTTGTCATTTGTCATTTGTCATTTG 1579
QY 769 PheValPheLeuLysValProGluThrLysGlyMetProLeuGluValIleThr---Glu 787
Db 1580 TTTCTTTTACACCGCTCTTGCCTTGCAGACCCCGGGAACCTCGAAGACATGGAAGGCT 1639
QY 788 PhePheAlaValGlyAlaLysGlnAlaAlaLysAla 800
Db 1640 TTTGGTACTTTTAGGTCCAAATCCACGCCAGCAAGGCT 1678

RESULT 8
US-09-291-922-21
; Sequence 21, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tinney, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 21
; LENGTH: 2017
; TYPE: DNA
; ORGANISM: Oryza sativa

US-09-291-922-21
Alignment Scores: 7.72e-42 Length: 2017
Pred. No.: 494.50 Matches: 185
Score: 35.71% Conservative: 101
Percent Similarity: 23.10% Mismatches: 206
Best Local Similarity: 12.13% Indels: 309
Query Match: 3 Gaps: 22
DB: 3
US-10-051-909-32 (1-800) x US-09-291-922-21 (1-2017)
QY 24 LeuProSerValValLeuAlaLeuProGlyProLeuProProAlaSerCysSerSerGln 43
Db 30 TTACTACTGACCGCCACTACTGTACACGGCCCA---GAGCGAGCCTCTCTCTCTCTGCA 86
QY 44 GluProValThrSerAspAspIleLeuGluAsp-----LysMetSerGly--- 58
Db 87 CCACCGAGATGGCTTCCGCGCGCTGCCGAGGCGCGCGCGCGGCGGCGGCGGCGGCGGCAAC 146
QY 59 -----AlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTyr 75
Db 147 GTCCGGTTCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTAC 206
QY 76 AspAsnAlaThrIleAlaAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsn 95
Db 207 GATATCGGGGTGATGAGCGGGCGTCTGCTGTACATCAAGAGGACTTCAACATC---AGT 263
QY 96 GluProThrValGlu-----GlyLeuIleValSerMetSerLeuIleGlyAlaThr 112
Db 264 GACGGAAGGTGGAGGTTCTCATGGGCATCTACTGCTCATCGGCTCCTTTC 323
QY 113 IleValThrThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIle 132
Db 324 GCG-----GCGGCGGAGCGTCCGACTGGATCGGCGGCGGCGGCGGCGGCGGCGGCGG 371
QY 133 LeuSerSerIleLeuTyrPhePheSerGlyLeuIleMetLeuTyrSerProAsnValTyr 152
Db 372 TTCGCGCGCGTCAATATCTTCGCGGGGGGTTCTCTCATGGGTTTCGCGGCGGCGGCGGCGG 431
QY 153 ValLeuLeuLeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuVal 172
Db 432 ATGCTCATGTTTCGCGCGCTTCTGTCGCGCATCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 491
QY 173 ProLeuTyrIleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuPro 192
Db 492 CCGGTGTACACCGCGGAGTGTTCGCGCGCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 551
QY 193 GlnPheSerGlySerGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeu 212
Db 552 GAGGTGTTCACTCAACTTCGCGCATCTGCTCGGTTAGCTCTCGAAGTATGCTTCTCTCCCGC 611
QY 213 SerPro---SerProAspTyrArgIleMetLeuGlyValLeuAlaIleProSerLeuPhe 231
Db 612 TTGCGCTGAACCTCGGCGTGGCGCATCATGCTCGGCGCATCGGCGGCGGCGGCGGCGGCGGCGG 668
QY 232 PhePheGlyLeuThrIlePheTyrLeuProGluSerProArgTyrLeuValSerLysGly 251
Db 669 CTGCTCGCGCTCATGGTGTCTGCGCATGCGGAGTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 728
QY 252 ArgMetAlaGluAlaLysLysValLeuGlnLysLeuArgGlyLysAspValSerGly 271
Db 729 CGCCTCGCGGACGCCAAGGTGCTGGAGAGACC----- 764
QY 272 GluLeuSerLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluTyr 291
Db 765 -----TCCGACACG----- 773
QY 292 IleIleGlyProAlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGln 311
Db 774 -----GCGGAGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 797
QY 312 IleThrLeuTyrGlyProGluGluGlyGlnSerTyrIleAlaArgProSerLysGlyPro 331

Db 797 ----- 797
QY 332 IleMetLeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSer 351
Db 797 ----- 797
QY 352 ValProLeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGln 371
Db 797 ----- 797
QY 372 AlaGlyGlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThr 391
Db 797 ----- 797
QY 392 AspGlnHisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspGluGlu 411
Db 797 ----- 797
QY 412 TyrAlaSerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSer 431
Db 797 ----- 797
QY 432 ArgGlnAlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAla 451
Db 797 ----- 797
QY 452 LeuSerMetArgArgGlnThrLeuLeuGlyGluGlyGlyAspGlyValSerSerThrAsp 471
Db 798 -----GCCGAC 803
QY 472 IleGlyGlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArg 491
Db 804 ATCAAGGCCGCC----- 815
QY 492 LysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArg 511
Db 816 -----GCCGGCATCCCTGAGGAGCTCGAC 839
QY 512 GlySerIleValSerLeuPro-----GlyGlyGlyAspValPheGluGlySerGluPhe 529
Db 840 GCGCAGCTGTGACCGTCCCAAGAGAGGAGCGGAAC----- 878
QY 530 ValHisAlaAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluPro 549
Db 879 -----GAGAAG 884
QY 550 ArgMetSerAspAlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrp 569
Db 885 CGGGTG-----TGG 893
QY 570 LysAspLeuPhe-----GluProGlyValArgArgAlaLeuLeuValGlyValGly 586
Db 894 AAGGAGTTCATCTGTCCCCGACCCCGGCATCGCGCATCTCTGTCCGGATCGGC 953
QY 587 IleGlnIleLeuGlnPheAlaGlyIleAsnGlyValLeuTyrThrProGlnIle 606
Db 954 ATCCACTTCTCCAGCATGCGTTGGGCATTCACCTCGTCTCTACAGCCCTCTCGTG 1013
QY 607 LeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAla 626
Db 1014 TTCAAGAGCCCCGGATTACG-----AACGACAACACATCTCTGGCACCACTTGGCCG 1067
QY 627 SerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMet 646
Db 1068 TTCGGTGTACCAAGAGGCTTTTCATCTCTTTG-----GCCACT 1106
QY 647 LeuLeuMetAspLeuSerGlyArgArgPheLeuLeuLeuGlyThrIleProIleLeuIle 666
Db 1107 TTCTTCATCGACGCGCTCGGCGCGCGCTGTGTGCTGGGACGACGCGGGGATAATC 1166
QY 667 AlaSerLeuVal-----IleLeuValValSerAsnLeuIleAspLeuGly 681

Db 1167 CTCTCCCTCATCGGCTCGGCGCGGGTCAACCGTCTGTCGGCCAGCACCCCGACGCAAG 1226
QY 682 ThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyrPheCysCysPheVal 701
Db 1227 ATACCTTGGGCCCATCGGCTTAAGCATCGCCTCCACCTCGCCTACGTCGCTTCTTCTCC 1286
QY 702 MetGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePheProThrArgValArg 721
Db 1287 ATCGGCTTGGCCCATCACGTTGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGC 1346
QY 722 -----GlyLeuCysIleAlaIleCysAlaPheThrPheTrpIleGlyAspIleVal 739
Db 1347 GCGCTGGGCTGCTCGCTCGGCTCGCGCCCAACCGCTCACGAGCGGCTCATCTCCATG 1406
QY 740 ThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyr 759
Db 1407 ACCTTCCTG-----TCGCTGTCCAAGGCCATCACCATCGGCGGAGCTTCTTCTCTAC 1460
QY 760 AlaValValCysLeuIleSerPheValPheValPheLeuLysValProGluThrLysGly 779
Db 1461 TCCGGCATCGCGGCTCGCTGGTGTCTTCTTCTACACCTACCTCCCGAGACCCGCGC 1520
QY 780 MetProLeuGluValIleThrGluPhePheAlaValGlyAlaLysGlnAlaAlaLys 799
Db 1521 CGGACGCTGGAGAGATGAGCAAGCTGTT-----GGCGACACGGCGCGCTCGGAA 1574
QY 800 Ala 800
Db 1575 TCA 1577
RESULT 9
US-09-291-922-19
; Sequence 19, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 19
; LENGTH: 1914
; TYPE: DNA
; ORGANISM: Zea mays
US-09-291-922-19
Alignment Scores:
Pred. No.: 1.46e-41 Length: 1914
Score: 491.50 Matches: 172
Percent Similarity: 35.36% Conservative: 107
Best Local Similarity: 21.80% Mismatches: 201
Query Match: 12.06% Indels: 309
DB: 3 Gaps: 18
US-10-051-909-32 (1-800) x US-09-291-922-19 (1-1914)
QY 41 SerSerGlnGluProValThr-----SerAspAspIle-----LeuGluAspLys 55
Db 35 TCAAGAAGTAGCCGTTAACGATGGCTTCCGAGGAGCTCGAAAGGCCGTCGAGCCAGG 94
QY 56 MetSerGly-----AlaValLeuValAlaIleValAlaSerIleGlyAsnLeu 71
Db 95 AAGAAGGGCAACGTCAGATATGCCTCCATATATGCTCCATCTGGCCTCCATGGCCTCTGTC 154
QY 72 LeuGlnGlyTrpAspAsnAlaThrIleAlaAlaValLeuTyrIleLysLysGluPhe 91

Db 155 ATCCTTGGCTATGACATTGGGGTGATGAGTGGAGCGGCCCATGTACATCAAGAGGACCTG 214
Qy 92 GlnLeuGlnAsnGluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAla 111
Db 215 AATATC---ACGGACGTGCAGCTGGAGATCCTGATCGGATCCTCAGTCTCTAC----- 265
Qy 112 ThrIleValThrThrPheSerGlyPro---LeuSerAspSerIleGlyArgArgProMet 130
Db 266 TCGCTGTTCGGATCCTTCGTGGCGCGCGGACGTCCGACAGGATCGGGCGCGCTTGACC 325
Qy 131 LeuIleLeuSerSerIleLeuTyrPhePheSerGlyLeuIleMetLeuTyrPheProAsn 150
Db 326 GTCGTGTTCGGCTGTCTCATCTCTTCGTGGGCTCGTTCATGGGTTTCGGCGTCAAC 385
Qy 151 ValTyrValLeuLeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThr 170
Db 386 TACGGCATGCTCATGGCGGCGCGCTCGTGGCCGAGTGGGTGGCTACGGGGGCATG 445
Qy 171 LeuValProLeuTyrIleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThr 190
Db 446 ATCGCGCCCGGTACACGGCCGAGATCTCGCCTCGCGGCTCCGCTGGCTTCCTGACCACC 505
Qy 191 LeuProGlnPheSerGlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMet 210
Db 506 TTCCCGGAGGTGTTTCATCAACATCGGCATCCTGTGGCTACCTGTCCAACCTCGCGTTC 565
Qy 211 SerLeuSerPro---SerProAspTyrArgIleMetLeuGlyValLeuAlaIleProSer 229
Db 566 GCGCGCTCCCGCTCCACCTCGGCTGGCGGTCTATGCTCGCCATTGGCGCAGTTCGCTCC 625
Qy 230 LeuPhePheGlyLeuThrIlePheTyrLeuProGluSerProArgTyrLeuValSer 249
Db 626 ---GGCCTGCTCGCGCTCCTGGTGTCTGTCATGCCCGAGTGCCTCGTGGCTGGTCTTG 682
Qy 250 LysGlyArgMetAlaGluAlaLysLysValLeuGlnLysLeuArgGlyLysAspVal 269
Db 683 AAGGGCGCCTCGCGGACCGCAGGCTGTGCTAGAGAAG----- 721
Qy 270 SerGlyGluLeuSerLeuLeuGluGlyLeuGluValGlyAspThrSerIleGlu 289
Db 721 ----- 721
Qy 290 GluTyrIleIleGlyProAlaThrGluAlaAspAspLeuValThrAspGlyAspLys 309
Db 721 ----- 721
Qy 310 GluGlnIleThrLeuTyrGlyProGluGluGlyGlnSerTrpIleAlaArgProSerLys 329
Db 722 -----ACCTTGCCACGCGACAGGAGCGCGCGCGCTGGCC----- 760
Qy 330 GlyProIleMetLeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsn 349
Db 760 ----- 760
Qy 350 GlnSerValProLeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMet 369
Db 760 ----- 760
Qy 370 ProGlnAlaGlyGlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSer 389
Db 760 ----- 760
Qy 390 ValThrAspGlnHisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAsp 409
Db 760 ----- 760
Qy 410 GluGluTyrAlaSerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeu 429
Db 760 ----- 760
Qy 430 LeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIleValHisGlyHisArgGly 449
Db 760 ----- 760

Qy 450 SerAlaLeuSerMetArgArgGlnThrLeuLeuGlyGluGlyGlyAspGlyValSerSer 469
Db 760 ----- 760
Qy 470 ThrAspIleGlyGlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsn 489
Db 761 ---GACATCAAGCGCGC----- 775
Qy 490 GlyArgLysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySer 509
Db 776 -----GCGGGGATTCCGAAGGCG 793
Qy 510 ArgArgGlySerIleValSerLeuPro-----GlyGlyGlyAspValPheGlu 525
Db 794 CTCGACGGGACGTAGTACCGTACCCGGCAAGGAGCAAGCGCGGTGAGTTG----- 847
Qy 526 GlySerGluPheValHisAlaAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGly 545
Db 847 ----- 847
Qy 546 LeuAlaGluProArgMetSerAspAlaAlaMetValHisProSerGluValAlaAlaLys 565
Db 847 ----- 847
Qy 566 GlySerArgTrpLysAspLeuPhe-----GluProGlyValArgArgAlaLeuLeu 582
Db 848 ---CAGGTGTGAAGAAGCTCATCTGTCCCCGACCCCGGCTGTCCGACGCATACTGCTC 904
Qy 583 ValGlyValGlyIleGlnIleLeuGlnPheAlaGlyIleAsnGlyValLeuTyrTyr 602
Db 905 TCGGCGCTGGGTCTCCACTTCTCCAGCAGGCTTCTGGCAGCGACTCCGTCGTCAGTAC 964
Qy 603 ThrProGlnIleLeuGluGlnAlaGlyValAlaIleLeuSerLysPheGlyLeuSer 622
Db 965 AGCGCGCGCTGTTCAGAGCGCGGGGATCACCGACGACAAAGCTCCTGGCGCTCACC 1024
Qy 623 SerAlaSerAlaSerIleLeuIleSerSerLeuThrLeuLeuMetLeuProCysIle 642
Db 1025 TGCGCG-----GTGGCGGTGACCAAGACGTTCCTTCATCCTG----- 1060
Qy 643 GlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuGlyThrIle 662
Db 1061 ---GTGGCCACGTTCCTGTGTGACCGCGGGCGGTGCGGCTCTGTGTGTATCAGCACG 1117
Qy 663 ProIleLeuIleAlaSerLeuValIleLeu-----ValValSerAsnLeu 677
Db 1118 GCGGGATGATTGTCTCGTCTCATCTGCTCGGGTGGGGCTCACCGTGGGGGCATCAC 1177
Qy 678 IleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyrPhe 697
Db 1178 CCGACACCAAGGTGCGGTGGCGGTGCGGCTGTGTCATCGGTCAACCTGTCTCTACATC 1237
Qy 698 CysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePhePro 717
Db 1238 GCCTTCTTCTCCATCGGCTCGGGCCCATCACCGGGGTGTACACCTCGGAAATATTCCCG 1297
Qy 718 ThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIleGlyAspIle 737
Db 1298 CTGAGGTGCGCGCTGGGCTTCGCGGTGGGTGGGTGGGAGCAACCGGTGTCACGCGCC 1357
Qy 738 IleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSer 757
Db 1358 GTCATCTCCATGACCTTCCTGTCCCTCTCCAAGGCCATCACCTCGGCGGCAGCTTCTTC 1417
Qy 758 IleTyrAlaValValCysLeuIleSerPheValPheValPheLeuLysValProGluThr 777
Db 1418 CTCTACTCCGGCATCGCGCGGTGCTTGGGTTTTCTTCTTCCGTCCCTCCCGGAGACA 1477
Qy 778 LysGlyMetProLeuGluValIleThrGluPhePhe-----Ala 790
Db 1478 CGCGCGGACGCTGGAGAGATGGGCAAGCTGTTCGGCATGCCAGACACGCGCATGGCT 1537

QY 791 ValGlyAlaLysGlnAlaAlaLys 799
| | | | |
Db 1538 GAAGAAGCAAGACGCCGAGCCAAG 1564

RESULT 10
US-09-291-922-27
; Sequence 27, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 27
; LENGTH: 1872
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-09-291-922-27

Alignment Scores:
Pred. No.: 2.59e-41 Length: 1872
Score: 489.00 Matches: 165
Percent Similarity: 34.22% Conservative: 106
Best Local Similarity: 20.83% Mismatches: 215
Query Match: 12.00% Indels: 306
DB: 3 Gaps: 14

US-10-051-909-32 (1-800) x US-09-291-922-27 (1-1872)

QY 24 LeuProSerValValLeuAlaLeuProGlyProLeuProProAlaSerCysSerSerGln 43
| | | | |
Db 119 ATGGCTTCTGTGGCTCCCGAGCGGGCGAGTCCATCCAGGAAC----- 166
| | | | |
QY 44 GluProValThrSerAspAspIleLeuGluAspLysMetSerGlyAlaValLeuValAla 63
| | | | |
Db 167 -----AAGGGCAATTTCAAGTACGCCTTCACCTGCGCC 199
| | | | |
QY 64 IleValAlaSerIleGlyAsnLeuLeuGlnGlyTrpAspAsnAlaThrIleAlaAlaAla 83
| | | | |
Db 200 CTCTGTGCTTCCATGGCCACCATCGTCTCGGCTACGACGTTGGGTGATGAGCGGTGG 259
| | | | |
QY 84 ValLeuTyIleLysLysGluPheGlnLeuGlnAsnGlu-----ProThrValGluGly 101
| | | | |
Db 260 TCGCTGTACATCAAGAGGGACCTGCAGATCAGGACGTGCAGTGGAGATCATGATGGGC 319
| | | | |
QY 102 LeuIleValSerMetSerLeuIleGlyAlaThrIleValThrPheSerGlyProLeu 121
| | | | |
Db 320 ATCCTGAGCGGTGTACGCGCTCATCGGGTCTCTCTC-----GGCGGAGGACG 367
| | | | |
QY 122 SerAspSerIleGlyArgArgProMetLeuIleLeuSerSerIleLeuTyIlePheSer 141
| | | | |
Db 368 TCCGACTGGGTGGCGCGCGGTACCGTGTCTTCGGCGCGCATCTTCAACAACGGC 427
| | | | |
QY 142 GlyLeuIleMetLeuTrpSerProAsnValTyIleValLeuLeuAlaArgPheValAsp 161
| | | | |
Db 428 TCCTTGCTCATGGGCTTCGCGTCAACTACGCCCATGTCTCATGTGGGCGCTTCGTACCG 487
| | | | |
QY 162 GlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyIleSerGluIleAlaPro 181
| | | | |
Db 488 GGAATCGGCGTGGGCTACGCCATCATGTGCGGCCAGTGTACACGCCCGGAGGTGTCCTCC 547
| | | | |
QY 182 SerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSerGlySerGlyMetPhe 201
| | | | |
Db 548 GCGTCGGCCCGCGGCTTCCTCACGTCTTTCACCGAGGTGTTTCATCATGTGGGCATCCTC 607
| | | | |

QY 202 LeuSerTyIleCysMetValPheGlyMetSerLeuSerPro---SerProAspTrpArgIle 220
| | | | |
Db 608 CTTGGCTACGTCTCAACTACGCCTTCGGCGCGCTCCCGCTCCACCTCAGCTGGCGCGTC 667
| | | | |
QY 221 MetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeuThrIlePheTyIleLeu 240
| | | | |
Db 668 ATGCTCGGCATCGCGCGCTCCCGTCCGCC---CTGCTTTCGCTCATGTGTTCGGCATG 724
| | | | |
QY 241 ProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGluAlaLysValLeu 260
| | | | |
Db 725 CCGGAGTCTCCTCGTGGCTCGTTCATGAAAGCGCGCTCGCGGACGCCAGGCGCGTTCGTG 784
| | | | |
QY 261 GlnLysLeuArgGlyLysAspAspValSerGlyGluLeuSerLeuLeuGluGlyLeu 280
| | | | |
Db 784 ----- 784
QY 281 GluValGlyGlyAspThrSerIleGluGluTyIleIleGlyProAlaThrGluAlaAla 300
| | | | |
Db 784 ----- 784
QY 301 AspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyIleGlyProGluGluGly 320
| | | | |
Db 784 ----- 784
QY 321 GlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGlySerValLeuSerLeu 340
| | | | |
Db 784 ----- 784
QY 341 AlaSerArgHisGlySerMetValAsnGlnSerValProLeuMetAspProIleValThr 360
| | | | |
Db 784 ----- 784
QY 361 LeuPheGlySerValHisGluAsnMetProGlnAlaGlyGlySerMetArgSerThrLeu 380
| | | | |
Db 784 ----- 784
QY 391 PheProAsnPheGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTrp 400
| | | | |
Db 795 -----GCCAAGACCTCCGACACG 802
| | | | |
QY 401 AspGluGluAsnLeuHisArgAspAspGluGluTyIleAlaSerAspGlyAlaGlyGlyAsp 420
| | | | |
Db 803 CCGGAGGAGCGCGTGGAGCGCCTTGACCAGATCAAGGCTGCC----- 844
| | | | |
QY 421 TyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLys 440
| | | | |
Db 844 ----- 844
QY 441 AspIleValHisHisGlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeu 460
| | | | |
Db 844 ----- 844
QY 461 GlyGluGlyGlyAspGlyValSerSerThrAspIleGlyGlyGlyTrpGlnLeuAlaTrp 480
| | | | |
Db 844 ----- 844
QY 481 LysTrpSerGluLysGluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyI 500
| | | | |
Db 844 ----- 844
QY 501 LeuHisGlnGluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGly 520
| | | | |
Db 845 -----GCCGCGATCCCTAGGGAACCTGACGGCGACGTTGGTTCGTTCATGCCTAAG--- 892
| | | | |
QY 521 GlyAspValPheGluGlySerGluPheValHisAlaAlaLeuValSerGlnSerAla 540
| | | | |
Db 892 ----- 892
QY 541 LeuPheSerLysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisProSer 560
| | | | |
Db 893 -----ACAAAGCGCGCCAGGAGAACGAGGTG----- 919
| | | | |
QY 561 GluValAlaAlaLysGlySerArgTrpLysAspLeuPhe-----GluProGlyVal 577
| | | | |

Db 920 -----TGAAGGAGCTCATCTTTTCGCCGACCCAGCCATG 955
Qy 578 ArgArgAlaLeuValGlyValGlyIleGlnIleLeuGlnPheAlaGlyIleAsn 597
Db 956 CGGCGCATACTGCTCGCGGCTCGGCATCCATTCTTCAGCAGGCGCGGCTCCGAC 1015
Qy 598 GlyValLeuTyrThrProGlnIleLeuGluGlnAlaGlyValAlaValIleLeuSer 617
Db 1016 TCCGTGCTGCTCTATAGCCACGCGTGTTCAGAGCGCGGCATCACCGGACAAACCAC 1075
Qy 618 LysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeu 637
Db 1076 CTGCTCGGC-----GCCACATGCGCCATGGGGGTTCATGAAGACGCTCTTC 1120
Qy 638 MetLeuProCysIleGlyPheAlaMetLeuLeuMetLeuSerGlyArgPheLeu 657
Db 1121 ATCTGT-----GTGCCCACGTTCCAGCTCGACCGCGTTCGGCAGCGCGCGCTG 1168
Qy 658 LeuLeuGlyThrIleProIleIleAlaSerLeuVal-----IleLeu 672
Db 1169 CTGCTGACACGACGCGCGCATGCTCGCCTGTCTCATCGGCCTCGGACGGGCTCAC 1228
Qy 673 ValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSer 692
Db 1229 GTCTGGGTGGCACCCCGGACCCCAAGTCCCGTGGGCCATCGGCCTGTGCATCGTGTCC 1288
Qy 693 ValIleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCys 712
Db 1289 ATCTGGCCTACGTGTCTTCTTCTCCATCGGCCTCGGCCCCCTCACCGAGGTGTACAC 1348
Qy 713 AlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPhe 732
Db 1349 TCGGAGGTCTTCCCACTGCGGTGCGCGCTGGGCTTCGGCTGGGCACGTCATGCAAC 1408
Qy 733 TrpIleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeu 752
Db 1409 CGGTGACACGCGCGGTGCTCCATGTCTTCTTCCTGTCTTGTCCAAAGGCCATCACCATC 1468
Qy 753 AlaGlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeu 772
Db 1469 GCGGCGAGCTTCTTCTGTACCGCGGCATCGCGCGCATAGGATGGATTTTCTTCTTCCAC 1528
Qy 773 LysValProGluThrLysGlyMetProLeuGluValIleThrGluPhePhe----- 789
Db 1529 TTCATTCGGAGACGCGTGGCGTGGCGCTCGAGGAGATAGGAAGCTTTTCGGCATGACG 1588
Qy 790 -----AlaValGlyAlaLysGlnAlaAlaLys 799
Db 1589 GACAGCGCGCTCGAAGCCCAAGACACCGCCACGAAA 1624

RESULT 11
US-09-291-922-25
; Sequence 25, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 25
; LENGTH: 2089
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-09-291-922-25

Alignment Scores: 1.58e-38 Length: 2089
Pred. No.: 463.50 Matches: 166
Score: 35.82% Conservative: 108
Best Local Similarity: 21.70% Mismatches: 200
Query Match: 11.37% Indels: 291
DB: 3 Gaps: 17
US-10-051-909-32 (1-800) x US-09-291-922-25 (1-2089)
Qy 52 LeuGluAspLysMetSerGly-----AlaValLeuValAlaIleValAlaSer 67
Db 104 GTCGAGCCCAAGAAAGAGGCAACGTGAGTTTCGCCCTTCGCCATCTCGCCTCC 163
Qy 68 IleGlyAsnLeuLeuGlnGlyTyrAspAsnAlaThrIleAlaAlaAlaValLeuTyrIle 87
Db 164 ATGACCTCCATCTCTCTCGCTACGACATCGGCGGTGATGAGCGGATCGCTGTATC 223
Qy 88 LysLysGluPheGlnLeuGlnAsnGluProThrValGlu-----GlyLeuIleVal 104
Db 224 CAGAGGATCTGAAGATC--AACGACACCCAGCTGAGGTCTCTATGGGCACTCTCAAC 280
Qy 105 SerMetSerLeuIleGlyAlaThrIleValThrPheSerGlyProLeuSerAspSer 124
Db 281 GTGFACTCGCTCATTTGGCTCTCTTCGCG-----GCGGGCGGACGCTCGACTGG 328
Qy 125 IleGlyArgArgProMetLeuIleLeuSerSerIleLeuTyrPhePheSerGlyLeuIle 144
Db 329 ATCGGCGGCGCTTACCATCGTCTTCGCGCGCGTCTTCTTCGCGGCGGCTCTCATC 388
Qy 145 MetLeuTrpSerProAsnValTyrValLeuLeuLeuAlaArgPheValAspGlyPheGly 164
Db 389 ATGGGCTTCTCCGTCAACTACGCCATGTCTCATGTTCGGCGCTTCGTGGCGGCGCATCGGC 448
Qy 165 IleGlyLeuAlaValThrLeuValProLeuTyrIleSerGluIleAlaProSerGluIle 184
Db 449 GTGGGTACGCTCTCATGTACGCGCGCGTGAACACGGCGGAGGTGTCCCGCGCTGTGCC 508
Qy 185 ArgGlyLeuLeuAsnThrLeuProGlnPheSerGlySerGlyMetPheLeuSerTyr 204
Db 509 CGTGGGTCTTCACATCTCTCCGAGGTGTTCATCAACTTCGGCATCTCTCTCGGATAT 568
Qy 205 CysMetValPheGlyMetSer--LeuSerProSerProAspTyrArgIleMetLeuGly 223
Db 569 GTCTCCAACCTTCGCCCTTCGCCCGCTCTCCCTCCGCTCGGCTCGGCTATGCTCGGC 628
Qy 224 ValLeuAlaIleProSerLeuPhePhePheGlyLeuThrIlePheTyrLeuProGluSer 243
Db 629 ATAGGCGGCTGCGGTCCGTC--CTGCTCGCGTTCATGGTGTCTCGGCATGCCGAGTCT 685
Qy 244 ProArgTrpLeuValSerLysGlyArgMetAlaGluAlaLysLysValLeuGlnLysLeu 263
Db 686 CCGCGGTGGCTCGTCAAGAGGCGCTCTCGCGACGCCAAGTTGTGCTTGCAAGACG 745
Qy 264 ArgGlyLysAspAspValSerGlyGluLeuSerLeuLeuGluGlyLeuValGly 283
Db 745 ----- 745
Qy 284 GlyAspThrSerIleGluGluTyrIleIleGlyProAlaThrGluAlaAlaAspLeu 303
Db 746 TCCGACACG----- 754
Qy 304 ValThrAspGlyAspLysGluGlnIleThrLeuTyrGlyProGluGluGlyGlnSerTrp 323
Db 755 -----CCGGAAGAGGCGCGCGGAGCGC 775
Qy 324 IleAlaArgProSerLysGlyProIleMetLeuGlySerValLeuSerLeuAlaSerArg 343
Db 776 ATCGCC----- 781
Qy 344 HisGlySerMetValAsnGlnSerValProLeuMetAspProIleValThrLeuPheGly 363

Db 781 ----- 781
QY 364 SerValHisGluAsnMetProGlnAlaGlyGlySerMetArgSerThrLeuPheProAsn 383
Db 781 ----- 781
QY 384 PheGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTrpAspGluGlu 403
Db 781 ----- 781
QY 404 AsnLeuHisArgAspAspGluGluTyrAlaSerAspGlyAlaGlyGlyAspTyrGluAsp 423
Db 782 ----- 799
424 AsnLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIleVal 443
Db 799 ----- 799
QY 444 HisHisGlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeuGlyGluGly 463
Db 799 ----- 799
QY 464 GlyAspGlyValSerSerThrAspIleGlyGlyGlyTrpGlnLeuAlaTrpLysTrpSer 483
Db 799 ----- 799
QY 484 GluLysGluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyrLeuHisGln 503
Db 799 ----- 799
QY 504 GluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGlyGlyAspVal 523
Db 800 ---GGCATCCCTCTGGGCCTCGACGGCGACGTGTTCCCGTGCCAAA----- 844
QY 524 PheGluGlySerGluPheValHisAlaAlaAlaLeuValSerGlnSerAlaLeuPheSer 543
Db 845 -----AAC 847
QY 544 LysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisProSerGluValAla 563
Db 848 AAAGGAAGCAGCGAGGAGAGCGCGTTTGAAGGACCTCATCTCTCACCACCATAGCC 907
QY 564 AlaLysGlySerArgTrpLysAspLeuPheGluProGlyValArgAlaLeuLeuVal 583
Db 908 -----ATGGCCACATCTCATCGCG 928
QY 584 GlyValGlyIleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyValLeuTyrThr 603
Db 929 GGAATCGGCATCCACTTCTTCCAGCAGTCTTCGGGCATCGACGCCCTCGTCTCTACAGC 988
QY 604 ProGlnIleLeuGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSer 623
Db 989 CCGTAGTTTTCAAGAGCGCGGCATCAG-----GGCGACAGCCGT 1030
QY 624 ---AlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIle 642
Db 1031 CTCCGGCGCACCCCGTGGCGGTTCGGGGCCACCAATACGGTCTTTCATCTG----- 1081
QY 643 GlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuGlyThrIle 662
Db 1082 ---GTGGCCACCTCCTCCTCGACCGCATCCGCGCGCGCGCTGCTGCTGACCGACG 1138
QY 663 ProIleLeuIleAlaSerLeuVal-----IleLeuValValSerAsnLeu 677
Db 1139 GGCGGCATGCTCGTCTCCTTAGTGGCCCTCGCGCGGGCTCACCGTTCATCAGCGCCAC 1198
QY 678 IleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyrPhe 697
Db 1199 CCGGACGAGAAGATCACCTGGGCCATCGTCTGTGCATCTTCTGCATCATGGCCTACGTG 1258
QY 698 CysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePhePro 717
Db 1259 GCCTTCTTCTCCATCGGCCTCGGCCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCG 1318

QY 718 ThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIleGlyAspIle 737
Db 1319 CTGCACGTGCGCGCTGGGTGCTCCCTGGCGGTGGCGTCAACCGCTGACCAGCGGC 1378
QY 738 IleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSer 757
Db 1379 GTGATCTCCATGACCTTCATTTGCTGTGCAAGGCCATGACCATCGCGCGCTTCTTC 1438
QY 758 IleTyrAlaValValCysLeuIleSerPheValPheValPheLeuLysValProGluThr 777
Db 1439 CTCTTCGCGGCATCGCTCATTCGCATGGGTGTTCTTCTTCGCTACCTGCGGAGACC 1498
QY 778 LysGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAla-----LysGln 795
Db 1499 CGCGCGCCGACGCTGGAGGACATGAGCTCGCTGTTTCGGCAACACGCGCACCAAGCAG 1558
QY 796 AlaAlaAlaLysAla 800
Db 1559 GCGCGCGCGGAGCC 1573

RESULT 12

US-09-489-039A-4762
; Sequence 4762, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 4762
; LENGTH: 1431
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-4762

Alignment Scores:

Pred. No.: 4.54e-37 Length: 1431
Score: 447.00 Matches: 158
Percent Similarity: 33.68% Conservative: 98
Best Local Similarity: 20.79% Mismatches: 186
Query Match: 10.97% Indels: 318
DB: 4 Gaps: 18

US-10-051-909-32 (1-800) x US-09-489-039A-4762 (1-1431)

QY 42 SerGlnGluProValThrSer---AspAspIleLeuGluAsp---LysMetSerGly--- 58
Db 4 TCGACAGATCAATAACTCAACTGGAGGCGGTATGCTGTGACACAAAAACAAGGCGT 63
QY 59 -----AlaValLeuValAlaIleValAlaSerIleGlyAsnLeuGln 73
Db 64 TCGAACAGACTATGACGTTCTTCGTCTGTTTCCTCGCGCGCTGGCTGGCTGCTGTT 123
QY 74 GlyTrpAspAsnAlaThrIleAlaAlaAlaValLeuTyrIleLysLysGluPheGlnLeu 93
Db 124 GGCCTTGATATCGGGTGTATTGCGGGTGCCCTTACCCCTTATTGCCAATGAGTCCAGATT 183
QY 94 GlnAsnGluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIle 113
Db 184 TCCGCCAC-----ACCCAGGAGTGGGTGTCAGTCCATGATGTTCCGGGGCTGCCGTC 237
QY 114 ValThrThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeu 133
Db 238 GCGCGCGTGGCAGCGGCTGGCTCTCTTTCAACTGGGCCGGAAGAGAGCTGATGATC 297
QY 134 SerSerIleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyrVal 153

Db 298 GGGCCCATCTCTTCGTCGGCGTTCGCTGTTCTCTGCGCGCGCGCCAAACGTCGAGATC 357
 QY 154 LeuLeuLeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValPro 173
 Db 358 CTGCTGGTTCCCGTGTGCTGCTCGGCTGGCGTGGCGTGGCGTGCCTCATATACGGCTCCG 417
 QY 174 LeuTyrIleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGln 193
 Db 418 CTGTATCTGTCGAAATCGCCCGAGAAAATTCGGCGAGTATGATTTCATGTACCAG 477
 QY 194 PheSerGlySerGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSer 213
 Db 478 CTGATCATCACCATCGGATCTTGGCGCTAT-----CTCTCTGACACCGCTTTCAGC 531
 QY 214 ProSerProAspTyrArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePhe 233
 Db 532 TACAGCGCGCATGCGCTGGATGCTCGGGTTATCATCATTCGGCGGTTTGTGCTG 591
 QY 234 GlyLeuThrIlePheTyrLeuProGluSerProArgTyrLeuValSerLysGlyArgMet 253
 Db 592 ---ATCGCGGTTATCTTCCTCGCGGACAGCCCGCGTGTGCGCGCAACGTCGCTTT 648
 QY 254 AlaGluAlaLysLysValLeuGlnLysLeuArgGlyLysAspValSerGlyGluLeu 273
 Db 649 GTCGATCGGAACGCGTGTCTGCGCCTGCGC-----GATACCGCGCGAA 681
 QY 274 SerLeuLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIle 293
 Db 682 -----GATACCGCGCGAA----- 696
 QY 294 GlyProAlaThrGluAlaAlaAspLeuValThrAspGlyAspLysGluGlnIleThr 313
 Db 696 ----- 696
 QY 314 LeuTyrGlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMet 333
 Db 696 ----- 696
 QY 334 LeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValPro 353
 Db 696 ----- 696
 QY 354 LeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly 373
 Db 696 ----- 696
 QY 374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393
 Db 696 ----- 696
 QY 394 HisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAla 413
 Db 697 ---GCGAAACGCGAGCTC----- 711
 QY 414 SerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGln 433
 Db 712 -----GATGAATCCGTGAAAGCCTGAAGTAAACAG 744
 QY 434 AlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeuSer 453
 Db 745 TCC----- 747
 QY 454 MetArgGlnThrLeuLeuGlyGluGlyGlyAspGlyValSerSerThrAspIleGly 473
 Db 747 ----- 747
 QY 474 GlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArgLysGlu 493
 Db 748 ---GGCTGGTCGCTG----- 759
 QY 494 GlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgGlySer 513
 Db 760 -----TTTAA----- 765

QY 514 IleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaAla 533
 Db 766 -----GACAAACAGCAACTTC----- 780
 QY 534 AlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAsp 553
 Db 780 ----- 780
 QY 554 AlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTyrLysAspLeuPhe 573
 Db 780 ----- 780
 QY 574 GluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnGlnPhe 593
 Db 781 -----CGCCGCGGGTGTTCCTCGGCATCCTGCTGAGGTGATGCAACAGTTC 828
 QY 594 AlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyValAla 613
 Db 829 ACCGGGATGAACGTATCATCTACTACGCGCCGAAGATCTTTGAGCTGGCGGTTATGCC 888
 QY 614 ValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeu 633
 Db 889 AACACCACTGAGCAATGTGG-----GGGACAGTATCGTCGCTGACT 933
 QY 634 ThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGly 653
 Db 934 AACGTGCTGGCCACCTTTATCGCCATCGGT-----CTGGTCGACCGCTGGGCG 981
 QY 654 ArgArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuVal 673
 Db 982 CGTAAA-----CCGACGCTGATCCTTGGCTTTATCGTATGGCC 1020
 QY 674 ValSerAsnLeuIleAspLeuGlyThrLeuAlaHis----- 685
 Db 1021 GCGGGAATGGCGTC---CTGGGTACCATCATGACATCGGCATTCACCTCTACCGCC 1077
 QY 686 AlaLeuLeuSerThrValSerValIleValTyrPheCysCysPheValMetGlyPheGly 705
 Db 1078 CAGTACATCGCCGTCCTGATGCTGATGATTCATCGTTCGCTTCGCCATGAGCGCGCC 1137
 QY 706 ProIleProAsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIle 725
 Db 1138 CCACTGATTGGGTACTGCTGCTCCGAAATCCAGCGCTGAAAGCGCGAGTTCGGTATC 1197
 QY 726 AlaIleCysAlaPheThrPheTrpIleGlyAspIleIleValThrTyrSerLeuProVal 745
 Db 1198 ACCTGCTCCACAGCGACCAACTGGATTGCCAATCATGATGTGCGCGCCACCTCTCTGACC 1257
 QY 746 MetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeuIle 765
 Db 1258 ATGCTCAACTCGCTGGCAGCGCAATACCTTCTGGGTGTACGGCGGTCTGAACGTGCTG 1317
 QY 766 SerPheValPheValPheLeuLysValProGluThrLysGlyMetProLeuGluValIle 785
 Db 1318 TTATCCTGCTGACGCTGTGGTGTATCCCGGAACCAAAACGCTCTCGCTGGAACATATT 1377

RESULT 13

US-09-489-039A-4731
 ; Sequence 4731, Application US/09489039A
 ; Patent No. 6610836
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary Breton et. al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
 ; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 2709.2004001
 ; CURRENT APPLICATION NUMBER: US/09/489,039A
 ; CURRENT FILING DATE: 2000-01-27
 ; PRIOR APPLICATION NUMBER: US 60/117,747
 ; PRIOR FILING DATE: 1999-01-29
 ; NUMBER OF SEQ ID NOS: 14342
 ; SEQ ID NO 4731
 ; LENGTH: 1545

```

; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-4731

Alignment Scores:
Pred. No.:      2.54e-36      Length:      1545
Score:          440.50      Matches:      164
Percent Similarity: 34.05%      Conservative: 105
Best Local Similarity: 20.76%      Mismatches: 206
Query Match:      10.81%      Indels:      315
DB:              4          Gaps:      16

US-10-051-909-32 (1-800) x US-09-489-039A-4731 (1-1545)

Qy 25 ProSerValValLeuAlaLeuProGlyProLeuProProAlaSer----- 39
Db 37 CCATATCTTGGCTTATCTCTGCTGACCCATTCCATCCACCAGTCACGTACTTTCGTCCTTA 96
Qy 40 -----CysSerSerGlnGlu-----ProValThrSerAspAspIleLeuGlu 53
Db 97 TGCTTACTCTGTGTGGCAGGAAAAAATGACTTCAATCAGTAACGACTCTACATTATCG 156
Qy 54 AspLysMetSerGlyAlaVal-----LeuValAlaIleValAlaSerIle 68
Db 157 CCGCGGACGCAACGTGATACCCGCGGATGAACCTGTTGTTCCATCGCCGCGCGGTG 216
Qy 69 GlyAsnLeuLeuGlnGlyTrpAspAsnAlaThrIleAlaAlaAlaValLeuTyrIleLys 88
Db 217 GCTGGCTGCTCTTTGGCCTGGATATCGGCGTGGTATCCGAGCGTTGCCCTTTATAACC 276
Qy 89 LysGluPheGlnLeuGlnAsnGluProThrValGluGlyLeuIleValSerMetSerLeu 108
Db 277 GACCAATTACCTTATCCAGCCAG-----CTTCAGGAGTGGTGGTTAGCAGCATGATG 330
Qy 109 IleGlyAlaThrIleValThrPheSerGlyProLeuSerAspSerIleGlyArgArg 128
Db 331 TTGGGGGCGCGATAGCGCGCTGTTAAACGGCTGGTGTTCGCTTCCGCTTGGCCGTAATA 390
Qy 129 ProMetLeuIleLeuSerSerIleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSer 148
Db 391 TACAGCCTGATGGCGGGCGGCTACTCTTGTGGCCGGCTCTATCGGATCCGCTTTTGCC 450
Qy 149 ProAsnValTyrValLeuLeuAlaArgPheValAspGlyPheGlyIleGlyLeuAla 168
Db 451 GCCAGCGTGGAGGTGCTGTGTGGCCGCGTGGTGTGGCGTGGCAGTCGGGATGGCC 510
Qy 169 ValThrLeuValProLeuTyrIleSerGluIleAlaProSerGluIleArgGlyLeuLeu 188
Db 511 TCTTATACCGCGCGCTGTACCTCTCCGAGATGGCCAGCAGAACGTCGCGGGAATG 570
Qy 189 AsnThrLeuProGlnPheSerGlySerGlyMetPheLeuSerTyrCysMetValPhe 208
Db 571 ATCAGTATGATCAGCTGATGGTCACCTTGGCATTGTGCTGCGGTTT-----CTTTCC 624
Qy 209 GlyMetSerLeuSerProSerProAspTrpArgIleMetLeuGlyValLeuAlaIlePro 228
Db 625 GATACCGCCTTAGCTACAGCGGTAACTGGCGGCCCATGCTGGGCGTGTGGCGCTGCCG 684
Qy 229 SerLeuPhePhePheGlyLeuThrIlePheTyrLeuProGluSerProArgTrpLeuVal 248
Db 685 GCGGTGATCCTGATCATCTTGTGCTCTTT---TTGCCGAACACGCCGCGCTGGCTGGCG 741
Qy 249 SerLysGlyArgMetAlaGluAlaLysLysValLeuGlnLysLeuArgGlyLysAspAsp 268
Db 742 GAGAGGGGACGCCATATCGAAGCGGAAGAGTGTGCGGATGCTGCGC----- 789
Qy 269 ValSerGlyGluLeuSerLeuLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIle 288
Db 790 -----GATACCTCG----- 798
Qy 289 GluGluTyrIleIleGlyProAlaThrGluAlaAlaAspAspLeuValThrAspGlyAsp 308
Db 799 -----GAAAAGCGCGCGCAGCTT----- 819
```

```

Qy 309 LysGluGlnIleThrLeuTyrGlyProGluGluGlnSerTrpIleAlaArgProSer 328
Db 819 ----- 819
Qy 329 LysGlyProIleMetLeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetVal 348
Db 820 -----AACGAGATC 828
Qy 349 AsnGlnSerValProLeuMetAspProIleValThrLeuPheGlySerValHisGluAsn 368
Db 829 CGTGAGAGCCTCAAGCTGAAGCAG----- 852
Qy 369 MetProGlnAlaGlyGlySerMetArgSerThrLeuPheProAsnPheGlySerMetPhe 388
Db 852 ----- 852
Qy 389 SerValThrAspGlnHisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAsp 408
Db 852 ----- 852
Qy 409 AspGluGluTyrAlaSerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerPro 428
Db 852 ----- 852
Qy 429 LeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIleValHisGlyHisArg 448
Db 852 ----- 852
Qy 449 GlySerAlaLeuSerMetArgArgGlnThrLeuLeuGlyGluGlyAspGlyValSer 468
Db 852 ----- 852
Qy 469 SerThrAspIleGlyGlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGlu 488
Db 853 -----GGCGGTTGG----- 861
Qy 489 AsnGlyArgLysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGly 508
Db 861 ----- 861
Qy 509 SerArgArgGlySerIleValSerLeuProGlyGlyGlyAspValPheGluGlySerGlu 528
Db 861 ----- 861
Qy 529 PheValHisAlaAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGlu 548
Db 862 -----GCGTTGTTTAAAG----- 873
Qy 549 ProArgMetSerAspAlaAlaMetValHisProSerGluValAlaAlaLysGlySerArg 568
Db 873 ----- 873
Qy 569 TrpLysAspLeuPheGluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGln 588
Db 874 -----ATCAATCGTAACGTGCGCGCGCGGTGTTCTCTGGSCATGCTGCTGCAG 921
Qy 589 IleLeuGlnGlnPheAlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGlu 608
Db 922 GCGATGCAGCAATTCCCGGCATGAACATCATCATGACTATGCGCCCGCGTATCTTTAAA 981
Qy 609 GlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIle 628
Db 982 ATGGCG-----GGCTTTACCCTACTGTAACAGCAGCATG 1014
Qy 629 LeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeu 648
Db 1015 ATCGCCACCCTGGTGGCGCTGACCTTTATGTTTGGCCACCTTTATTGCGGTGTTTACA 1074
Qy 649 MetAspLeuSerGlyArgArg---PheLeuLeuLeuGlyThrIleProIleLeuIleAla 667
Db 1075 GTGGATAAAGCGGCGCGCAAGCCAGCGCTGAAATCGGCTTTAGCGTGATGGCGCTGGGC 1134
```


QY 299 AlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyrGlyProGlu 318
Db 793 GCG-----CAGAAGGAG----- 804
QY 319 GluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGlySerValLeu 338
Db 804 ----- 804
QY 339 SerLeuAlaSerArgHisGlySerMetValAsnGlnSerValProLeuMetAspProIle 358
Db 805 -----TACGATGACATC 816
QY 359 ValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGlyGlySerMetArgSer 378
Db 817 TCTACCTGATC----- 828
QY 379 ThrLeuPheProAsnPheGlySerMetPheSerValThrAspGlnHisAlaLysAsnGlu 398
Db 829 -----AAATTGAA 837
QY 399 GlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAlaSerAspGlyAlaGly 418
Db 838 -----GCTGGT 843
QY 419 GlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGlu 438
Db 844 AATAAGTAC-----ACGCCCCAGAGCACCTTTGCG--- 873
QY 439 GlyLysAspIleValHisHisGlyHisArgGlySerAlaLeuSerMetArgGlnThr 458
Db 873 ----- 873
QY 459 LeuLeuGlyGluGlyGlyAspGlyValSerSerThrAspIleGlyGlyGlyTrpGlnLeu 478
Db 873 ----- 873
QY 479 AlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArgLysGluGlyGlyPheLysArg 498
Db 873 ----- 873
QY 499 ValTyrLeuHisGlnGluGlyValProGlySerArgGlySerIleValSerLeuPro 518
Db 874 -----ACGATTGTCAAAACACCC 891
QY 519 GlyGlyGlyAspValPheGluGlySerGluPheValHisAlaAlaAlaLeuValSerGln 538
Db 891 ----- 891
QY 539 SerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHis 558
Db 891 ----- 891
QY 559 ProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuPheGluProGlyValArg 578
Db 892 -----TGG-----ATCCTC 900
QY 579 ArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnGlnPheAlaGlyIleAsnGly 598
Db 901 AAAATTTTGTGTCGGCATTACCTGGGCGGCACCTGCAGCAACACCACCGGTTAACGTC 960
QY 599 ValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyValAlaValIleLeuSerLys 618
Db 961 ATTATGTATTACGGTACGGAAATCCTGAGTCCCGCGTTTC----- 1002
QY 619 PheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMet 638
Db 1003 -----TCAGAGCGTACTTCGCTGATCTGCAACGTCTGAACGGTGTCTCTCC 1050
QY 639 Leu-----ProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArg--- 655
Db 1051 GTGGGCGGTATGCTGATCGGG---GTACTGTTCCTTGTCCACCGCTTCAAACGTAAGACA 1107
QY 656 -----PheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuVal 670

Db 1108 ATCATCATCTACGGTTTCGCCATTATGCAACCCCTGCATCTGATTATCGCCGCTGTCGAT 1167
QY 671 IleLeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThr 690
Db 1168 TACACCCCTGTTGGT-----GACCTGAAAGCTACCGGCATCTGGCTGCTGGGCGCG 1218
QY 691 ValSerValIleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIle 710
Db 1219 CTGTTCTGTCGGGTGATGCAGGGCTCC-----ATGGGCTTT-----ATTACCTGGGTG 1266
QY 711 LeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPhe 730
Db 1267 GTATTAGCGGAACCTGTTCCCGCTGAAATTCGCGGACTGTGCGATGGGATCTCGGTATTT 1326
QY 731 ThrPheTrpIleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIle 750
Db 1327 TTCAATGTTGATCATGAACCGCGTGGTCAAGTATCTGTTCCCGCTACTGCAGGCAAACTG 1386
QY 751 GlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheVal 770
Db 1387 GGCCTTGGGCCAGTCTTCTTTATCTTCCGCCCATTAATACTATTAGCCATTCTTTTCGTG 1446
QY 771 PheLeuLysValProGluThrLysGlyMetProLeuGlu 783
Db 1447 GTCTTTGCCCTGCCCGAGACCTCCAATAAATCGCTTGAG 1485

Search completed: October 13, 2004, 23:30:06
Job time : 239.106 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
OM protein - nucleic search, using frame_plus_p2n model
Run on: October 13, 2004, 16:55:37 ; Search time 1141.17 Seconds
(without alignments)
3566.233 Million cell updates/sec

Title: US-10-051-909-32
Perfect score: 4075
Sequence: 1 IRSGSWLAVQTPFTPDLDRLR.....PLEVITEFFAVGAKQAAKA 800

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3372238 seqs, 2543544081 residues
Total number of hits satisfying chosen parameters: 6744476

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool_h/US10051909/runat_13102004_123339_20010/app_query.fasta_1.1678
-DB=Published_Applications NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10051909@cgn_1_1_953@runat_13102004_123339_20010
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:
11: /cgn2_6/ptodata/1/pubpna/US09E_PUBCOMB.seq:
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:
18: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:
19: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:
20: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result * Query

No.	Score	Match	Length	DB	ID	Description
1	4075	100.0	2777	13	US-10-051-909-31	Sequence 31, Appl
2	4059	99.6	2908	15	US-10-310-154-350	Sequence 350, App
3	3517	86.3	2824	13	US-10-051-902-1	Sequence 1, Appli
4	3517	86.3	2824	13	US-10-051-909-1	Sequence 1, Appli
5	2869	70.4	2031	16	US-10-425-114-32961	Sequence 32961, A
c	2835.5	69.6	2769	17	US-10-437-963-93472	Sequence 93472, A
	2674	65.6	2601	13	US-10-051-902-7	Sequence 7, Appli
8	2674	65.6	2601	13	US-10-051-909-7	Sequence 7, Appli
9	2584	63.4	2190	9	US-09-938-842A-1315	Sequence 1315, Ap
10	2584	63.4	2190	11	US-09-938-842A-1315	Sequence 1315, Ap
11	2571	63.1	2671	16	US-10-424-599-10324	Sequence 10324, A
12	2543	62.4	2405	16	US-10-425-114-7958	Sequence 7958, Ap
13	2483.5	60.9	2573	16	US-10-424-599-108435	Sequence 108435,
14	2442.5	59.9	2271	16	US-10-425-114-29146	Sequence 29146, A
15	2275	55.8	2205	9	US-09-938-842A-2254	Sequence 2254, Ap
16	2275	55.8	2205	11	US-09-938-842A-2254	Sequence 2254, Ap
17	2245.5	55.1	2432	17	US-10-437-963-47424	Sequence 47424, A
18	1963	48.2	1959	17	US-10-437-963-92959	Sequence 92959, A
c	1763	43.3	2231	17	US-10-437-963-70909	Sequence 70909, A
	1670.5	41.0	1692	13	US-10-051-902-9	Sequence 9, Appli
20	1670.5	41.0	1692	13	US-10-051-909-9	Sequence 9, Appli
21	1670.5	41.0	1692	13	US-10-051-909-9	Sequence 9, Appli
22	1545	37.9	3205	16	US-10-424-599-77967	Sequence 77967, A
23	1424	34.9	1487	13	US-10-051-902-13	Sequence 13, Appl
24	1424	34.9	1487	13	US-10-051-909-13	Sequence 13, Appl
25	1392	34.2	1313	17	US-10-767-701-12595	Sequence 12595, A
26	1252.5	30.7	1289	16	US-10-424-599-42000	Sequence 42000, A
27	1035	25.4	1412	16	US-10-260-238-747	Sequence 747, App
c	949.5	23.3	1259	16	US-10-424-599-101136	Sequence 101136,
	943	23.1	1009	13	US-10-051-902-15	Sequence 15, Appl
29	943	23.1	1009	13	US-10-051-909-15	Sequence 15, Appl
30	943	23.1	1009	13	US-10-051-909-15	Sequence 15, Appl
31	813	20.0	761	16	US-10-425-114-35910	Sequence 35910, A
32	793	19.5	650	16	US-10-425-114-3864	Sequence 3864, Ap
33	663	16.3	435	17	US-10-767-701-18822	Sequence 18822, A
34	620	15.2	711	17	US-10-767-701-12614	Sequence 12614, A
35	620	15.2	751	16	US-10-260-238-2683	Sequence 2683, Ap
36	615.5	15.1	2253	17	US-10-437-963-49329	Sequence 49329, A
37	604	14.8	674	17	US-10-437-963-60088	Sequence 60088, A
38	600	14.7	778	16	US-10-260-238-2684	Sequence 2684, Ap
39	587.5	14.4	1374	15	US-10-369-493-47011	Sequence 47011, A
40	580	14.2	583	17	US-10-021-323-11734	Sequence 11734, A
41	562	13.8	870	13	US-10-051-902-5	Sequence 5, Appli
42	562	13.8	870	13	US-10-051-909-5	Sequence 5, Appli
43	554	13.6	627	16	US-10-425-114-12008	Sequence 12008, A
44	552	13.5	659	16	US-10-260-238-2685	Sequence 2685, Ap
45	549	13.5	1898	16	US-10-424-599-96711	Sequence 96711, A

ALIGNMENTS

RESULT 1
US-10-051-909-31
; Sequence 31, Application US/10051909
; Publication No. US20020199217A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Helentjaris, Tim
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB1163 US CIP
; CURRENT APPLICATION NUMBER: US/10/051,909
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 60/083,044
; PRIOR FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 31
; LENGTH: 2777
; TYPE: DNA
; ORGANISM: Zea mays

Db 1992 ACAATCCAAATCTTGATAGCATCTCTAGTTATCCCTGGTTGTGTGCCAATCTAAATGATTTG 2051
Qy 681 GlyThrLeuAlaHisAlaLeuSerThrValSerValIleValTyrPheCysPhe 700
Db 2052 GGTACACTAGCCCATGCTTTGCTCTCCACCGTCAGTGTATCGTCTACTTCTGCTGCTTC 2111
Qy 701 ValMetGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePheProThrArgVal 720
Db 2112 GTTATGGGATTTGGTCCCATCCCAACATTTTATGTGCAGAGATCTTTCCAAACCGGTT 2171
Qy 721 ArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIleGlyAspIleIleValThr 740
Db 2172 CGTGGCCTCTGTATTGCCATTTTGCCCTTTACATTTCTGGATCGGAGATATCATCGTCACC 2231
Qy 741 TyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleIleTyrAla 760
Db 2232 TACAGCCTTCCTGTGATGCTGATGCTATTTGGACTGGCGGTGTTTTCAGCATATATGCA 2291
Qy 761 ValValCysLeuIleSerPheValPheValPheLeuLysValProGluThrLysGlyMet 780
Db 2292 GTCGTATGCTTGATTTCTTTTGTTGTTGCTCTTCTTAAGTCCCTGAGACAAAGGGATG 2351
Qy 781 ProLeuGluValIleThrGluPhePheAlaValGlyAlaLysGlnAlaAlaLysAla 800
Db 2352 CCCCTGAGGTTATTACCGAATTTCTTGCAATTGTTGCGAAGCAAGCGGCTGCCAAAGCC 2411

RESULT 2

US-10-310-154-350
; Sequence 350, Application US/10310154
; Publication No. US20030233670A1

GENERAL INFORMATION:

; APPLICANT: Edgerton, Michael D
; APPLICANT: Chomet, Paul S.
; APPLICANT: Adams, Thomas H.
; APPLICANT: Ruff, Thomas G.
; APPLICANT: Agarwal, Ameeta K.
; APPLICANT: Ahrens, Jeffrey E.
; APPLICANT: Ball, James A.
; APPLICANT: Banu, G.
; APPLICANT: Bell, Erin
; APPLICANT: Boddupalli, Raghava
; APPLICANT: Deikman, Jill
; APPLICANT: Deng, Molian
; APPLICANT: Dong, Jinzhao
; APPLICANT: Duff, Stephen M.
; APPLICANT: Galligan, Meghan M.
; APPLICANT: Hinchey, Brenda S.
; APPLICANT: Huang, Shihshieh
; APPLICANT: Johnson, G. Richard
; APPLICANT: Jung, Vincent
; APPLICANT: Kretzmer, Keith A
; APPLICANT: Laccetti, Lucille B.
; APPLICANT: Lai, Chao-Qiang
; APPLICANT: Lee, Gary
; APPLICANT: Lin, Jie-Yi
; APPLICANT: Liu, Jingdong
; APPLICANT: Lu, Bin
; APPLICANT: Luethy, Michael M.
; APPLICANT: Lund, Adrian
; APPLICANT: Madson, Linda L.
; APPLICANT: Malloy, Kathleen A.
; APPLICANT: McKiel, Christine L.
; APPLICANT: Miller, Philip W.
; APPLICANT: Padmavathi, Manchikanti
; APPLICANT: Parnell, Laurence D.
; APPLICANT: Start, William G.
; APPLICANT: Tennesen, Dan
; APPLICANT: Vidya, K.R.
; APPLICANT: Wang, Haiyun
; APPLICANT: Xin, Zhanguo
; APPLICANT: Xu, Nanfei
; APPLICANT: Yang, Chunzhi
; APPLICANT: Zeng, Xiaoping

; APPLICANT: Zhang, Qiang
; APPLICANT: Zhao, Yajuan
; APPLICANT: Zhou, Li
; TITLE OF INVENTION: Gene Sequences and Uses Thereof in Plants
; FILE REFERENCE: 38-15(52796)B
; CURRENT APPLICATION NUMBER: US/10/310,154
; CURRENT FILING DATE: 2002-12-04
; PRIOR APPLICATION NUMBER: 60/337,358
; PRIOR FILING DATE: 2001-12-04
; NUMBER OF SEQ ID NOS: 736
; SEQ ID NO 350
; LENGTH: 2908
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (331)..(2565)
; OTHER INFORMATION:
US-10-310-154-350

Alignment Scores:

Pred. No.: 0 Length: 2908
Score: 4059.00 Matches: 797
Percent Similarity: 99.75% Conservative: 1
Best Local Similarity: 99.62% Mismatches: 2
Query Match: 99.61% Indels: 0
DB: 15 Gaps: 0

US-10-051-909-32 (1-800) x US-10-310-154-350 (1-2908)

Qy 1 IleArgSerGlySerTrpLeuAlaValGlnThrProPheThrProAspLeuAspArgArg 20
Db 166 ATTCGGAGCGGCTCTTGGCTTGAGTCCAGACGCCCTTCACCCCTGATCTGGACCGGAG 225
Qy 21 GluArgLeuLeuProSerValValLeuAlaLeuProGlyProLeuProProAlaSerCys 40
Db 226 GAGCGGCTCCTTCGTCAGTTGTTCTTGGCTTGGCTGGCTGGCTCTTCCGCTGCTTCGTGT 285
Qy 41 SerSerGlnGluProValThrSerAspAspIleLeuGluAspLysMetSerGlyAlaVal 60
Db 286 TCTTCACAGGAGCGGTGACCTCGGACGATATCTTGGAGGACAAGATGTCGGGGGCTGT 345
Qy 61 LeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrpAspAsnAlaThrIle 80
Db 346 CTTGTGCCCATAGTCGCTCCATCGGCAATCTATTGACGGGTGGGACAATGCCACCATC 405
Qy 81 AlaAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsnGluProThrValGlu 100
Db 406 GCAGTGTGTTCTGTATATAAAGAGGAATTTCAATTGCAAAATGAGCCCACTGTGGAG 465
Qy 101 GlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThrPheSerGlyPro 120
Db 466 GGACTAATTGTCAATGTACATTATCGGCGCCACCATCGTTACTACATTCTCCGGGCCA 525
Qy 121 LeuSerAspSerIleGlyArgArgProMetIleLeuSerSerIleLeuTyrPhePhe 140
Db 526 TTATCAGACTCGATTGGCGGACGCCCTATGCTTATCTCTTCAATTCTGTACTTCTTC 585
Qy 141 SerGlyLeuIleMetLeuTrpSerProAsnValTyrValLeuLeuAlaArgPheVal 160
Db 586 AGCGGCTCATCATGCTATGGTCTCCTAATGCTATGCTATGCTGCTGTGGCAGCGCTTCGTA 645
Qy 161 AspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyrIleSerGluIleAla 180
Db 646 GATGGATTGGTATGGCTTGGCTGTCAACGCTGTGCTTGTACATTTTCAGAAATAGCC 705
Qy 181 ProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSerGlySerGlyMet 200
Db 706 CCTTCGGAGATTAGAGGTTTGTGAATACACTACCACAATTTCAGTGGATCAGGAGGATG 765
Qy 201 PheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSerProAspTrpArgIle 220
Db 766 TTCTTGTCACTACTGCATGGTGTGTTGGGATGTCCTGTGCGCCATCACCCGATGGAGAATT 825

Qy 221 MetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeuThrIlePheTyrLeu 240
Db 826 ATGCTTGGTGTCTCGCGATACCTTTCATTGTTCTTCTTGGTTTGACAAATATTATCTT 885
Qy 241 ProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGluAlaLysLysValLeu 260
Db 886 CCTGAATCTCCAAGATGGCTCGTTAGCAAAAGGTCGATGGCAGAGGCAAAAAGGTTG 945
Qy 261 GlnLysLeuArgGlyLysAspValSerGlyGluLeuSerLeuLeuGluGlyLeu 280
Db 946 CAAAAGTTACGGGGAAAGACCATGTCACAGTGAAGTTCAGGTGAATGTCCTTCTTCGAAGGTTG 1005
Qy 281 GluValGlyGlyAspThrSerIleGluGluTyrIleIleGlyProAlaThrGluAlaAla 300
Db 1006 GAGTTGGAGGAGACACTTCCATTTGAAGAGTACATCATTTGGACCTGCCACCGAGCGACCC 1065
Qy 301 AspAspLeuValThrAspGlyAspLysGluGluIleThrLeuTyrGlyProGluGluGly 320
Db 1066 GATGATCTTGTACTGACGGTGATAAGGAACAAATCACACTTTATGGGCTTGAAGAAGGC 1125
Qy 321 GlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGlySerValLeuSerLeu 340
Db 1126 CAGTCATGATGTCGACCTTCCAAAGGACCCAGCATGCTTGAAGTGTCTTCTCTT 1185
Qy 341 AlaSerArgHisGlySerMetValAsnGlnSerValProLeuMetAspProIleValThr 360
Db 1186 GCATCTCGTCATGGGAGCATGGTGAACCCAGAGTGTACCCCTTATGGATCCGATTGTGACA 1245
Qy 361 LeuPheGlySerValHisGluAsnMetProGlnAlaGlyGlySerMetArgSerThrLeu 380
Db 1246 CTTTTTGGTAGTGTCCATGAGAATATGCCTCAAGCTGAGGAAGTATGAGGAGCACATTG 1305
Qy 381 PheProAsnPheGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTrp 400
Db 1306 TTTCCAAACTTTGGAAGTATGTTTCAGTGTCCACAGATCAGCATGCCAAATGAGCAGTGG 1365
Qy 401 AspGluGluAsnLeuHisArgAspAspGluGluTyrAlaSerAspGlyValAlaGlyLys 420
Db 1366 GATGAAGAGAAATCTTTCATAGGATGACGAGGATGACGATCTGATGGTGCAGGAGGTGAC 1425
Qy 421 TyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAlaThrGlyValAlaGluGlyLys 440
Db 1426 TATGAGGACAAATCTCCATAGCCCATTTGCTGTCCAGGCAGGCAACAGGTGCGGAAGGAAG 1485
Qy 441 AspIleValHisHisGlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeu 460
Db 1486 GACATTGTGCACCATGTTACCGTGGAAAGTGTCTTGAGCATGAGAAAGGCAAGCCCTCTTA 1545
Qy 461 GlyGluGlyGlyAspGlyValSerSerThrAspIleGlyGlyGlyTrpGlnLeuAlaTrp 480
Db 1546 GGGGAGGTGGAGATGTTGTGAGCAGCACTGATATCGTGGGGGATGGCAGCTTGTCTTGG 1605
Qy 481 LysTrpSerGluLysGluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyr 500
Db 1606 AAATGGTTCAGAGAAGGAAGGTGAGAAATGGTAGAAAGGAAGGTGGTTTCAAAAGAGTCTAC 1665
Qy 501 LeuHisGlnGluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGly 520
Db 1666 TTGCACCAAGAGGGAGTTCCTGGCTCAAGAGGGGCTCAATTGTTTCACTTCCCGGTGGT 1725
Qy 521 GlyAspValPheGluGlySerGluPheValHisAlaAlaLeuValSerGlnSerAla 540
Db 1726 GCGGATGTTCTTGAGGGTAGTGAGTTTGTACATGCTGCTGCTTGTAGTAAGTCAGTCAGCA 1785
Qy 541 LeuPheSerLysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisProSer 560
Db 1786 CTTTCTCAAAGGGTCTGCTGAACCCAGCATGTCAGATGCTGCCATGCTTCAACCATCT 1845
Qy 561 GluValAlaAlaLysGlySerArgTrpLysAspLeuPheGluProGlyValArgAla 580
Db 1846 GAGGTAGCTGCCAAAGGTTACCGTTGGAAGAATTGTTTGAACCTCGAGTCAGGCGTGCC 1905

Qy 581 LeuLeuValGlyValGlyIleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyValLeu 600
Db 1906 CTGTAGTCCGGTGTGGAAATCAGATCCTTCAACAGTTTGTGGAATAAACCGTGTCTG 1965
Qy 601 TyrTyrThrProGlnIleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGly 620
Db 1966 TACTATACCCCAAAATTTCTTGAGCAAGCTGGTGGCAGTTATTCTTTCCAAATTTGGT 2025
Qy 621 LeuSerSerAlaSerAlaSerIleLeuIleSerSerSerLeuThrThrLeuLeuMetLeuPro 640
Db 2026 CTCAGCTCGGCATCAGCATCCATCTTGTATCAGTTCTCTCACTACCTTACTAATGCTTCT 2085
Qy 641 CysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuGly 660
Db 2086 TGCATTGGCTTTGCCATGCTGCTTATGGATCTTTCCGGAAGAAGGTTTTTGTGCTAGGC 2145
Qy 661 ThrIleProIleLeuIleAlaSerLeuValIleLeuValSerAsnLeuIleAspLeu 680
Db 2146 ACAATTCCAATCTTGATAGCATCTCTAGTTATCTCTGTTGTGTCCAATCTAATGATTG 2205
Qy 681 GlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyrPheCysCysPhe 700
Db 2206 GGTACACTAGCCCATGCTTGTCTCTCCACCGTCAGTGTATCGTCTACTTCTGCTGCTTC 2265
Qy 701 ValMetGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePheProThrArgVal 720
Db 2266 GTTATGGGATTTGGTCCCATCCCCAACATTTATGTGCAGAGATCTTTTCCAAACCGGTT 2325
Qy 721 ArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIleGlyAspIleIleValThr 740
Db 2326 CGTGGCCTCTGTATTGGCATTTTGTGCTTACATTTGGATCGGAGATATCATCGTCACC 2385
Qy 741 TyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAla 760
Db 2386 TACAGCCTTCTGTGATGCTGAATGCTATTGGACTGGCGGTGTTTTTTCAGCATATATGCA 2445
Qy 761 ValValCysLeuIleSerPheValPheValPheLeuLysValProGluThrLysGlyMet 780
Db 2446 GTCGTATGCTTGATTTCTTTGTTGTTCTTCTTAAAGTCCCTGAGACAAAGGGGATG 2505
Qy 781 ProLeuGluValIleThrGluPhePheAlaValGlyAlaLysGlnAlaAlaLysAla 800
Db 2506 CCCCTTGAGGTTATTACCGAATTTCTTTGCAAGTGGTGGGAAGCAAGCGGCTGCAAAAGCC 2565

RESULT 3
US-10-051-902-1
; Sequence 1, Application US/10051902
; Publication No. US20020178468A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/10/051,902
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-14
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 2824
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (29)
; NAME/KEY: unsure
; LOCATION: (622)
; NAME/KEY: unsure
; LOCATION: (636)
; NAME/KEY: unsure


```
; LOCATION: (638)
; NAME/KEY: unsure
; LOCATION: (669)
; NAME/KEY: unsure
; LOCATION: (771)
; NAME/KEY: unsure
; LOCATION: (822)
; NAME/KEY: unsure
; LOCATION: (856)
; NAME/KEY: unsure
; NAME/KEY: unsure
; LOCATION: (889)
; NAME/KEY: unsure
; LOCATION: (896)
; NAME/KEY: unsure
; LOCATION: (944)
US-10-051-902-1
Alignment Scores:
Pred. No.: 0
Score: 3517.00
Percent Similarity: 95.18%
Best Local Similarity: 91.57%
Query Match: 86.31%
DB: 13
US-10-051-909-32 (1-800) x US-10-051-902-1 (1-2824)
QY 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrp 75
Db 238 ATGGGGGCGCGTGTATGGTCGCCATCGCGCCTCTATCGCAACTTGTGCAGGGCTGG 297
QY 76 AspAsnAlaThrIleAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsn 95
Db 298 GACAAATCGACAAATTGCTGGAGCCCTCTGTACATAAAGAAGGAATTCAACCTGCAGAGC 357
QY 96 GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115
Db 358 GAGCCTCTGATCGAAGGCCTCATCGTCGCCATGTTCTCTCATGGGCAACAGTCATACA 417
QY 116 ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer 135
Db 418 ACATCTCGGGGCCAAGGGCTGACTCGGTGGTAGGAGGCCCATGCTGGTCGCCCTCGGCT 477
QY 136 IleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyrValLeuLeu 155
Db 478 GTCCTCTACTTCGTCACTGGGTGGTGATGCTTTGGGCGCCAAATTGTGTACATCTTGCTC 537
QY 156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr 175
Db 538 CTCGCAAGGCTCATTGATGGGTTTCGTATCGGTTTGGCGGTGCACACTTGTTCCTCTCTAC 597
QY 176 IleSerGluIleAlaProSerGluIleArg---GlyLeuLeuAsnThrLeuProGlnPhe 194
Db 598 ATCTCCGAAACTGCACCGCACAGANATTCTTGGGGCTGNTNGAACACGTTTGGCGCAGTTC 657
QY 195 SerGly---SerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSer 213
Db 658 ATTGGGGTCAGNGGAGGATGTTCTCTCTCTACTGTCATGGTGTGGGATGTCCCTCATG 717
QY 214 ProSerProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePhe 233
Db 718 CCCAAACCTGATTGGAGGCTCATGCTTGGAGTTCTGTGCATCCCGTCACTTATNACTTT 777
QY 234 GlyLeuThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMet 253
Db 778 GGACTGACTGTCTTCTACTTGCCTGAATCACCAAGGTGGCTGTGNAGCAAAGGAAGGATG 837
QY 254 AlaGluAlaLysLysValLeuGlnLysLeuArgGlyLysAspAspValSerGlyGluLeu 273
Db 838 GCGGAGGCGAAGAGAGTNTGCAAGGCTGCGGGGAAGAGAGATGCTCTCANGGGAGANG 897
QY 274 SerLeuLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIle 293
:::|||||:::|||||
```

```
Db 898 GCTCTTCTAGTTGAAGGTTTGGGGGTCGGTAAAGATACACGTAATTNAGAGTACATCATT 957
QY 294 GlyProAlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThr 313
|||
Db 958 GGACCTGCCACCGAGGCAGCCGATGATCTTTGTAACCTGACGGTGATAAGGAACAAATCACA 1017
314 LeuTyrGlyProGluGluGlnSerTrpIleAlaArgProSerLysGlyProIleMet 333
|||
Db 1018 CTTTATGGCCTGAAGAAGGCCAGTCATGGATTGCTCGACCTTCTAAGGGACCCATCATG 1077
334 LeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValPro 353
|||
Db 1078 CTTTGAAGTGTGCTTCTCTTGCATCTCGTCATGGGAGCATGGTGAACACAGAGTGTACCC 1137
354 LeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly 373
|||
Db 1138 CTTATGGATCCGATTGTGACACTTTTGGTAGTGTCCATGAGATATGCCTCAAGCTGGA 1197
374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393
|||
Db 1198 GGAAGTATGAGGAGCACATTGTTTCCAAACTTTTGAAGTATGTTGAGTGTACAGATCAG 1257
394 HisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAla 413
|||
Db 1258 CATGCCAAAAATGACGAGTGGGATGAAGAGAATCTTCATAGGGATGACGAGGAGTACGCA 1317
414 SerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGln 433
|||
Db 1318 TCTGATGGTCAGGAGGTGACTATGAGGACAATCTCCATAGCCCATTTGCTTCCAGGCAG 1377
434 AlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeuSer 453
|||
Db 1378 GCAACAGGTGCCGAAGGAAGGACATTGTGCACCATGGTCCACCGTGAAGTGCTTTGAGC 1437
454 MetArgArgGlnThrLeuLeuGlyGluGlyAspGlyValSerSerThrAspIleGly 473
|||
Db 1438 ATGAGAAAGGCAAAAGCCTCTTAGGGGAGGTGGAGATGGTGTGAGCAGCAGTATCGGT 1497
474 GlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArgLysGlu 493
|||
Db 1498 GGGGGATGGCAGCTTGTCTTGGAAATGGTCAGAGAAGGAAGGTGAGATGGTAGAAGGAA 1557
494 GlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArgGlySer 513
|||
Db 1558 GGTGTTTCAAAGAGTCTACTTGCCAAAGAGGGAGTTCCTGGCTCAAGAAAGGGCTCA 1617
514 IleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaAla 533
|||
Db 1618 ATTGTTTCACTTCCCGTGGTGGCGATGTTCTTGAGGGTAGTGAGTTGTACATGCTGCT 1677
534 AlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAsp 553
|||
Db 1678 GCTTTAGTAAGTCAGTCAGCACTTTTCTCAAAGGGTCTTGTCTGAACCCACGATGTAGAT 1737
554 AlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuPhe 573
|||
Db 1738 GCTGCCATGGTTCACCCATCTGAGGTAGCTGCCAAAGGTTACGTTGGAAGATTGTTT 1797
574 GluProGlyValArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnGlnPhe 593
|||
Db 1798 GAACCTGGAGTGAGGCGTGCCCTGTTTACTCGGTGTTGGAATTCAGATCCTTCAACAGTTT 1857
594 AlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyValAla 613
|||
Db 1858 GCTGGAATAAACGGTGTCTGTACTATACCCCAACAAATTTCTTGAGCAAGCTGGTGGCA 1917
614 ValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeu 633
|||
Db 1918 GTTATTCTTCCAAATTTGGTCTCAGCTCGGCATCAGCATCCATCTTGATCAGTTCTCTC 1977
634 ThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGly 653
|||
Db 1978 ACTACCTTACTAATGCTTCTTGTGATTGGCTTGGCCATGCTGCTTATGGATCTTTCCGGA 2037
```


QY 334 LeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValPro 353
Db 1078 CTTGGAAAGTGCTTTCTCTGTCATCTCGTATGGAGCATGGTGAACACAGAGTGACCC 1137
QY 354 LeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly 373
Db 1138 CTTATGATCCGATTGTGACACTTTTGGTAGTGTCCATGAGAAATATGCCTCAAGCTGA 1197
QY 374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393
Db 1198 GGAAGATGAGGAGCACATTGTTCCAACTTTGGAAAGTATGTTTCAAGTGTACAGATCAG 1257
QY 394 HisAlaLysAsnGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAla 413
Db 1258 CATGCCAAAATGACAGTGGGATGAAGAGAAATCTTTCATAGGATGACGAGGATACGCA 1317
QY 414 SerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGln 433
Db 1318 TCTGATGGTGACGAGGTGACTATGAGGACAATCTCCATAGCCCATTTGCTGTCCAGGCAG 1377
QY 434 AlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeuSer 453
Db 1378 GCAACAGGTGCGGAAGGGAAGGACATTGTGCACCATGTTGACCGTGGAAAGTGCTTTGAGC 1437
QY 454 MetArgArgGlnThrLeuLeuGlyGluGlyGlyAspGlyValSerSerThrAspIleGly 473
Db 1438 ATGAGAAGGCAAAAGCCTCTTAGCGAGGGTGGAGATGGTGTGAGCAGCACATGATACGGT 1497
QY 474 GlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArgLysGlu 493
Db 1498 GGGGGATGGCAGCTTGTCTTGGAAATGGTCAGAGAAGGAGGTGAGAAATGGTAGAAGGAA 1557
QY 494 GlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArgGlySer 513
Db 1558 GGTGGTTTCAAAAGAGTCTACTTGCACCAAGAGGGAGTTCTGGCTCAAGAAGGGGCTCA 1617
QY 514 IleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaAla 533
Db 1618 ATTGTTTCACTTCCCGGTGGTGGCGATGTTCTTGAGGGTAGTGAGTTTGTACATGTGCT 1677
QY 534 AlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAsp 553
Db 1678 GCTTTAGTAAGTCAGTCAGCACATTCTCAAAGGGTCTTGTGTAACCAACGATGTGAGAT 1737
QY 554 AlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuphe 573
Db 1738 GCTGCCATGTTTCAACCATCTGAGGTAGTGTGCCAAAGGTTACGTTGGAAAGATTGTTT 1797
QY 574 GluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnGlnPhe 593
Db 1798 GAACCTGGAGTGAGGGCTGCCCTGTGTAGTCGGTGTGGAAATTCAGATCCTTCAACAGTTT 1857
QY 594 AlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyValAla 613
Db 1858 GCTGGAATAAACGGTGTCTGTACTATACCCCAAAATTTCTTGAGCAAGCTGGTGTGGCA 1917
QY 614 ValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeu 633
Db 1918 GTTATTCTTCCAAATTTGGTCTCAGCTCGGCATCAGCATCCATCTTGTATCAGTTCTCTC 1977
QY 634 ThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGly 653
Db 1978 ACTACCTTACTAATGCTTCTTGTGATGGCTTTGCCATGCTGCTTATGGATCTTTCCGGA 2037
QY 654 ArgArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuVal 673
Db 2038 AGAAGGTTTTTGTGCTAGGCACAAATCCAAATCTTGTATAGCATCTCTAGTTATCCTGGTT 2097
QY 674 ValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerVal 693
Db 2098 GTGTCCAATCTAATTGATTGGGTACACTAGCCCATGCTTTGCTCTCCACCATCAGTGTT 2157
QY 694 IleValTyrPheCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAla 713

Db 2158 ATCGTCTACTTCTGCTGCTTGGTATGGGATTTGGTCCCATCCCAACATTTTATGTGCA 2217
QY 714 GluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrp 733
Db 2218 GAGATCTTTCCAAACCAGGGTTCTGGCCTCTGTATTGCCATTGTGCCTTTACATTTCTGG 2277
QY 734 IleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAla 753
Db 2278 ATCGGAGATATCATCTGCCTACCTACAGCCTTCTGTGATGCTGATGCTATTGGACTGGCG 2337
QY 754 GlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeuLys 773
Db 2338 GGTGTTTTTACGACATATATGCAGTCGTATGCTTGATTTCTTTGTGTTCTGCTTCTTAAAG 2397
QY 774 ValProGluThrLysGlyMetProLeuGluValIleThrGluPheAlaValGlyAla 793
Db 2398 GTCCCTGAGACAAAGGGGATGCCCTTGAGGTTATTACCGAATTTCTTTCAGATTGGTGCG 2457
QY 794 LysGlnAlaAlaLysAla 800
Db 2458 AAGCAAGCGGCTGCAAAAGCC 2478
RESULT 5
US-10-425-114-32961
; Sequence 32961, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingsong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 32961
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLMO17045C09_FLI
US-10-425-114-32961

Alignment Scores:
Pred. No.: 2,09e-274 Length: 2031
Score: 2869.00 Matches: 561
Percent Similarity: 99.65% Conservative: 1
Best Local Similarity: 99.47% Mismatches: 2
Query Match: 70.40% Indels: 0
DB: 16 Gaps: 0

US-10-051-909-32 (1-800) x US-10-425-114-32961 (1-2031)
QY 237 IlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGluAla 256
Db 2 ATATTTTATCTCTCGAATCTCCAAGATGGCTCGTTAGCAAAGTCGGATGGCAGGCA 61
QY 257 LysLysValLeuGlnLysLeuArgGlyLysAspValSerGlyGluLeuSerLeuLeu 276
Db 62 AAAAAGGTGTTCCAAAAGTTACGGGGGAAAGACGATGTCTCAGGTGAATTGTCCCTTCTT 121
QY 277 LeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIleGlyProAla 296
Db 122 CTCGAAGGGTTGGAGGATGGAGGAGACACTTCCATTGAAGAGTACATCATTTGGACCTGCC 181
QY 297 ThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyrGly 316
Db 182 ACCGAGGCAGCGCATGATCTTGTGTACTGACGGTGATAAGGAACAATCACACTTTATGGG 241

Qy 317 ProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGlySer 336
Db 242 CTGAAGAAGCCAGTCATGGATTGCTCGACCTTCCAAGGGACCCAGCATGCTTGGAGT 301
Qy 337 ValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValProLeuMetAsp 356
Db 302 GTGCTTTCTCTTGCATCTCGTCATGGGAGCATGGTGAACAGAGTGTACCCCTTATGGAT 361
Qy 357 ProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGlyGlySerMet 376
Db 362 CCGATTGTGACACTTTTGGTAGTGTCCATGAGAAATATGCCTCAAGCTGGAGGAAGTATG 421
Qy 377 ArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGlnHisAlaLys 396
Db 422 AGGAGCACATTGTTTCCAAACTTTTGGAAAGTATGTTCAAGTGTACAGATCAGCATGCCAAA 481
Qy 397 AsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTrpAlaSerAspGly 416
Db 482 AATGAGCAGTGGGATGAAGAGAATCTTTCATAGGGATGACGAGGAGTACGCATCTGATGGT 541
Qy 417 AlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAlaThrGly 436
Db 542 GCAGGAGGTGACTATGAGGACAATCTCCATAGCCCATGCTGTCCAGGAGGCAACAGGT 601
Qy 437 AlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeuSerMetArgArg 456
Db 602 GCGGAAGGGAAGGACATTGTGCACCATGTTCAACGTGGAAGTGTCTTTCAGCATGAGAAG 661
Qy 457 GlnThrLeuLeuGlyGluGlyGlyAspGlyValSerSerThrAspIleGlyGlyTrp 476
Db 662 CAAAGCCTCTTAGGGGAGGTGGAGATGTTGTGAGCAGCACTGATATCGTGGGGATGG 721
Qy 477 GlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArgLysGluGlyPhe 496
Db 722 CAGCTTGCTTGGAAATGGTCAGAGAAGGAAGGTGAGATGGTAGAAGGAAGGTGTTTC 781
Qy 497 LysArgValTyrLeuHisGlnGluGlyValProGlySerArgArgGlySerIleValSer 516
Db 782 AAAAGAGTCTACTTGCACCAAGAGGGAGTCTTGGCTCAAGAGGGGTCAATTGTTTCA 841
Qy 517 LeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaAlaAlaLeuVal 536
Db 842 CTTCCCGGTGGTGGCGATGTTCTTGAGGGTAGTAGTGTGATGCTGCTGCTTTAGTA 901
Qy 537 SerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAspAlaAlaMet 556
Db 902 AGTCAGTCAGCACCTTTCTCAAAGGGTCTTGTGTAACCAACGATGTCAGATGCTGCCATG 961
Qy 557 ValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuPheGluProGly 576
Db 962 GTTCACCCATCTGAGGTAGCTGCCAAAGGTTTCAAGTTGGAAGATTGTTTGAACCTGGA 1021
Qy 577 ValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnGlnPheAlaGlyIle 596
Db 1022 GTGAGGCGTGCCTGTAGTGGTGTGGAAATTCAGATCCTTCAACAGTTTGTCTGGAATA 1081
Qy 597 AsnGlyValLeuTyrTrpThrProGlnIleLeuGluGlnAlaGlyValAlaValIleLeu 616
Db 1082 AACGGTGTCTGTACTATACCCCAACAAATCTTGAGCAAGCTGGTGTGGCAGTTATTCTT 1141
Qy 617 SerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThrLeu 636
Db 1142 TCCAAATTTGGTCTCAGCTCGGCATCAGCATCCATCTTGTATCAGTTCTCTCACTACCTTA 1201
Qy 637 LeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPhe 656
Db 1202 CTAATGCTTCTTGCATTGGCTTTGCCCATGCTGCTTATGGATCTTTCGGAAGAAGGTTT 1261
Qy 657 LeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuValValSerAsn 676
Db 1262 TTGCTGCTAGGCACAAATCCAAATCTTGATAGCATCTCTAGTTATCCTTGGTGTGTCCAAT 1321

Qy 677 LeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyr 696
Db 1322 CTAATTGATTGGGTACACTAGCCCATGCTTTGCTCTCCACCGTCAGTGTATCGTCTAC 1381
Qy 697 PheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePhe 716
Db 1382 TTCTGTGCTTTCGTTATGGGATTTGGTCCCATCCCAACATTTTATGTGAGAGATCTTT 1441
Qy 717 ProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIleGlyAsp 736
Db 1442 CCAACCAAGGTTTCGTGGCTCTGTATTGCCATTGTGCTTACATTCTGGATCGGAGAT 1501
Qy 737 IleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPhe 756
Db 1502 ATCATCGTCACCTACAGCTTCTGTGTGATGTGAATGCTATTGGACTGGCGGTGTTTC 1561
Qy 757 SerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeuLysValProGlu 776
Db 1562 ACATATATGCAGTCGTATGCTTGTATTTCCTTTGTGTTCTTCTTAAGGTCCCTGAG 1621
Qy 777 ThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAlaLysGlnAla 796
Db 1622 ACAAGGGGATGCCCCCTTGAGTTATTACCGAATTCTTTGCAGTTGGTGGCAAGCAAGCG 1681
Qy 797 AlaAlaLysAla 800
Db 1682 GCTGCAAAAGCC 1693

RESULT 6

US-10-437-963-93472/c
; Sequence 93472, Application US/10437963
; Publication No. US20040123343A1

GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 93472

; LENGTH: 2769

; TYPE: DNA

; ORGANISM: Oryza sativa

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT4530_91853C.1

US-10-437-963-93472

Alignment Scores:

Pred. No.: 7.17e-271 Length: 2769
Score: 2835.50 Matches: 562
Percent Similarity: 82.45% Conservative: 77
Best Local Similarity: 72.52% Mismatches: 120
Query Match: 69.58% Indels: 16
DB: 17 Gaps: 7

US-10-051-909-32 (1-800) x US-10-437-963-93472 (1-2769)

Qy 30 AlaLeuProGlyProLeuProPro-AlaSerCys-----SerSerG1 43
Db 2684 GCCCTTCCAGATCCAGCCGCTCTCTTCTTTAGGGGATCCGAAATCTCGGTGGACG 2625
Qy 43 nGluProValThrSerAspAspIleLeuGluAspLysMetSerGlyAlaValLeuValA1 63
Db 2624 AGAGACTTGGTGGTAAGAT-----TCGCCGCCCATGGCGGCGCGTGTGTCGC 2574

QY 63 aIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrpAspAsnAlaThrIleAlaAlaI 83
Dbb 2573 CATCGCGGCTCCATCGCAACTTGTCTGAGGGCTGGATAATGCAACCATTCAGAGTGC 2514
QY 83 aValLeuTyrIleLysLysGluPheGlnLeuGlnAsnGluProThrValGluGlyLeuI 103
Dbb 2513 GGTACTGTACATCAAGAAGGAATTCAACTTGCAGAGCGGCCCTTATCGAAGGCCTGAT 2454
QY 103 eValSerMetSerLeuIleGlyAlaThrIleValThrThrPheSerGlyProLeuSerAs 123
Dbb 2453 CGTGGCCATGTCTGCTCATTTGGGCGGACGATCATCAGCAGCTTCTCTGGAGCAGTGGCTGA 2394
QY 123 pSerIleGlyArgArgProMetLeuIleLeuSerSerIleLeuTyrPhePheSerGlyLe 143
Dbb 2393 TTCTTTTGGTAGGCGGCCCATGTGATCGCTCGCTCGCTGCTCTACTTTTGTAGTGGGCT 2334
QY 143 uIleMetLeuTrpSerProAsnValTyrValLeuLeuLeuAlaArgPheValAspGlyPh 163
Dbb 2333 AGTGATGCTTTGGCGGCCAAATGTGTATGTGTCTCTGGCAGGCTCAITGACGGGTT 2274
QY 163 eGlyIleGlyLeuAlaValThrLeuValProLeuTyrIleSerGluIleAlaProSerGl 183
Dbb 2273 CGGGATCGGTTTGGCTGTTCACGCTGTACCAATTGTATCATCTCTGAGACTGCCCGACGGA 2214
QY 183 uIleArgGlyLeuLeuAsnThrLeuProGlnPheSerGlySerGlyGlyMetPheLeuSe 203
Dbb 2213 CATCAGAGGACTGCTAAACACCGCTGCCGCAGTTTCAGTGGGTCTGGAGGGATGTTCTTTC 2154
QY 203 rTyrCysMetValPheGlyMetSerLeuSerProSerProAspTrpArgIleMetLeuGl 223
Dbb 2153 ATACTGCATGGTATTTGGCATGTCTCTCATGCCACAGCCAGATTGGAGGATCATGCTTGG 2094
QY 223 yValLeuAlaIleProSerLeuPhePheGlyLeuThrIlePheTyrLeuProGluSe 243
Dbb 2093 CGTTCTATCAATACCATCATCTATATACTTTTGCATTGACCATCTTTTACTTACCITGAATC 2034
QY 243 rProArgTrpLeuValSerLysGlyArgMetAlaGluAlaLysLysValLeuGlnLysLe 263
Dbb 2033 GCCGAGGTGGCTCGTGAGCAAAAGGAAGATGGCTGAGGCCAACGCTGTGTGTGCAAGGCCT 1974
QY 263 uArgGlyLysAspAspValSerGlyGluLeuSerLeuLeuGluGlyLeuGluValGl 283
Dbb 1973 GCGTGAAGAGAGAAGATGTTTCAGGAGAAATGGCCCTTCTCGTTGAAGGTCTGGGGGTGG 1914
QY 283 yGlyAspThrSerIleGluGluTyrIleIleGlyProAlaThrGluAlaAlaAspAspLe 303
Dbb 1913 GAAAGACACAAAAATTGAGGAATACATAATTGGACCTGATGATGAGCTTGCTGATGAAGG 1854
QY 303 uValThrAspGlyAspLysGluGlnIleThrLeuTyrGlyProGluGluGlyGlnSerTr 323
Dbb 1853 GCTGGCTCCA--GATCCAGAGAAGATCAAACTGTATGGTCCCTGAAGAAGCCTTATCGTG 1797
QY 323 pIleAlaArgProSerLysGlyProIleMetLeuGlySerValLeuSerLeuAlaSerAr 343
Dbb 1796 GGTGGCCCGTCTCTTTCACGGGCAAAAGTGCACCTTGAAGTGCATTAGTCTCATCTCTCG 1737
QY 343 gHisGlySerMetValAsnGlnSerValProLeuMetAspProIleValThrLeuPheGl 363
Dbb 1736 TCATGGTAGTATGTCAGTCAGGGTAAGCCCTTGTGGATCCCTGTGTCAACCTTTTGG 1677
QY 363 ySerValHisGluAsnMetProGlnAlaGlyGlySerMetArgSerThrLeuPheProAs 383
Dbb 1676 AAGTGTCCATGAGAAGATGCCCTGAGATTAATGGGAAGCATGCGGAGCACATTGTTTCCCTAA 1617
QY 383 nPheGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTrpAspGluGl 403
Dbb 1616 CTTTGGCAGCATGTTTAGTGTGGCGGAACACAGCAGCAAGCTAAAGGTGATTGGGATGCTGA 1557
QY 403 uAsnLeuHisArgAspAspGluGluTyrAlaSerAspGlyAlaGlyGlyAspTyrGluAs 423
Dbb 1556 GAGT---CAACGGGAGGGTGAAGATTATGGATCAGACCATTGGTGGGATGACATTGAAGA 1500
QY 423 pAsnLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIle-- 442

Dbb 1499 TAGCCTCCAAAAGCCCACTTATTCTCTCGTCAAGCGACAAGCGTGAAGGAGATCGC 1440
QY 443 -ValHisHisGlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeuGlyGl 462
Dbb 1439 TGCACCTCATGGCAGTATAATGGTGTCTGGGA-----AGAAAGTAGTAGTCTCATGCA 1386
QY 462 uGlyGlyAspGlyValSerSerThrAspIleGlyGlyGlyTrpGlnLeuAlaTrpLysTr 482
Dbb 1385 GGGCGGGAGGCAGTAAGCAGCATGGGCATTGGTGGGGATGGCAGTTGGCTTGGAATG 1326
QY 482 pSerGluLysGluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyrLeuHi 502
Dbb 1325 GACTGAGAGAGAAGGTGCAGATGGCGAAAAAGAGGTGGCTTCCAACGTATCTACTTGCA 1266
QY 502 sGlnGluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGlyGlyAs 522
Dbb 1265 TGAAGAGGTGTGACAGGTGATCGCAGGGGCTCTATACTGTCTATGCTTGCCT---GGAGGTGA 1209
QY 522 pValPheGluGlySerGluPheValHisAlaAlaAlaLeuValSerGlnSerAlaLeuPh 542
Dbb 1208 TGTTCCTCCTGGTGGTGGTTCGTCAGGCGAGTGTCTTGTGTCAGCAACCTGCTCTTTA 1149
QY 542 eSerLysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisProSerGluVa 562
Dbb 1148 CTCTAAGGAATTGATGGAGCAACGCCCTTGCTGGCCCTGCTATGGTGCATCCATCTCAGGC 1089
QY 562 lAlaAlaLysGlySerArgTrpLysAspLeuPheGluProGlyValArgArgAlaLeuLe 582
Dbb 1088 AGTTGCTAAAGTCCAAAATGGGCAGACTTATTGGAACCTGGAGTGAAGCATGCTCTGTT 1029
QY 582 uValGlyValGlyIleGlnIleLeuGlnPheAlaGlyIleAsnGlyValLeuTyrTy 602
Dbb 1028 TGTGGCATAGGATACAAATCCTGCAACAGTTTGTGGCATTATGGAGTTCTGTACTA 969
QY 602 rThrProGlnIleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSe 622
Dbb 968 CACTCCACAAATCTTGAGCAAGCTGGTGTGTTCTTCTTGCAAAACATTGGACTTAG 909
QY 622 rSerAlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIl 642
Dbb 908 CTCCTCATCTGCATCTATTCTTATTAGCGGACTGACAACCTTGTCTGATGCTTCCCAGCAT 849
QY 642 eGlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuLeuGlyThrIl 662
Dbb 848 TGGTATTGCTATGAGGCTCATGGATATGCTGGAAGAGGTTTCTTCTCTCTTGCACAAAT 789
QY 662 eProIleLeuIleAlaSerLeuValIleLeuValValSerAsnLeuIleAspLeuGlyTh 682
Dbb 788 CCCTATCCTGATAGTACACTAGCTATCTGATCTGGTCAATATTCTGGATGTGGGAC 729
QY 682 rLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyrPheCysCysPheValMe 702
Dbb 728 CATGGTTTCATGCCCTCACTGTCCACAGTCAGTCACTACTCTACTTCTGCTTCTTGTCTAT 669
QY 702 tGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePheProThrArgValArgGl 722
Dbb 668 GGGGTTGCGGCCCTATTCCAAACATTCTCTGTGCAGAGATTTTCCCGACCCCGTTCGTGG 609
QY 722 yLeuCysIleAlaIleCysAlaPheThrPheTrpIleGlyAspIleIleValThrTyrSe 742
Dbb 608 CATCTGCATAGCCCATCTGTGCCCTAACATTCTGGATCGGTGATATCAATTGTGACATACAC 549
QY 742 rLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValVa 762
Dbb 548 CCTCCCCGTGATGCTCAACGCCCATTTGGACTCGCTGGAGTGTTTGGAAATCTACGCAGTGGT 489
QY 762 lCysLeuIleSerPheValPheValPheLeuLysValProGluThrLysGlyMetProLe 782
Dbb 488 CTGCATACTGGCTTTCTGCTTTTGTCTTTCATGAAGGTGCCGAGACAAAGGGCATGCCTCT 429
QY 782 uGluValIleThrGluPhePheAlaValGlyAlaLysGlnAla 796

Db 428 TGAAGTCATCACCAGGTTCTTCTCTGTGCGAGCAAGCAGGCC 386

RESULT 7

US-10-051-902-7
; Sequence 7, Application US/10051902
; Publication No. US20020178468A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/10/051,902
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-14
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 7
; LENGTH: 2601
; TYPE: DNA
; ORGANISM: Glycine max
US-10-051-902-7

Alignment Scores:

Pred. No.:	7.15e-255	Length:	2601
Score:	2674.00	Matches:	522
Percent Similarity:	82.20%	Conservative:	92
Best Local Similarity:	69.88%	Mismatches:	115
Query Match:	65.62%	Indels:	18
DB:	13	Gaps:	8

US-10-051-909-32 (1-800) x US-10-051-902-7 (1-2601)

QY	56	MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrp	75
Db	175	ATGAAAGGTGCGCTCCTTGTGCTATTGCGCGCTTCCATTGGTAATTTCTCCAAGGATGG	234
QY	76	AspAsnAlaThrIleAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsn	95
Db	235	GATAATGCTACCATCGCCGGGGCTAATGGTTACATTAAGAAAGACCTTGCTTTGGGAACA	294
QY	96	GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr	115
Db	295	-----ACTATGGAAGGCTTGTGTGGCATGTCCCTGATTGGAGCAACGGTAATCACC	348
QY	116	ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer	135
Db	349	ACATGCTCTGGTCTCTATAGCGGATTGGCTCGGTCGGCGACCCATGATGATAATCTCATCT	408
QY	136	IleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyrValLeuLeu	155
Db	409	GTGCTCTATTCTTGGTGGTTTGGTGATGCTGTGGTCCCCAAATGTGTATGTTGTGC	468
QY	156	LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr	175
Db	469	TTGGCGAGGCTACTTGTATGGATTGGGATTGGCTTGTGTGACTCTGTCCCGGTCTAT	528
QY	176	IleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSer	195
Db	529	ATATCTGAACCGCGCGCTCTGAAATAAGGGGTCGTTGAATACGCTTCCCTCAGTTTCACT	588
QY	196	GlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSer	215
Db	589	GGCTCTGGAGGAATGTTTGTCTACTGTATGGTTTGTGGCATGTCTTGTAGTCCCGCG	648
QY	216	ProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeu	235
Db	649	CCTAGCTGGAGGCTCATGCTTGGGTTCTGTCTATTCCTTCTCTCTCTCTTGTGTCATTG	708
QY	236	ThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGlu	255

Db	709	ACCATTTTTTCTTCCCGAGTCTCCTCGGTGGTGGTCAGCAAGGAGGATGCTCGAG	768
QY	256	AlaLysLysValLeuGlnLysLeuArgGlyLysAspValSerGlyGluLeuSerLeu	275
Db	769	GCTAAGAAGGTGCTCCAAAGATTGCGCGAAGGGAGGATGTGTGTCAGCGGAGATGGCATTG	828
QY	276	LeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIleGlyPro	295
Db	829	CTGGTTGAAGGTCTCGGGATTGGGGTGATACATCTATCGAAGAGTACATAATTGGCCCT	888
QY	296	AlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyr	315
Db	889	GCTGACGATGTGGCTGATGTCATGAACATGCAACAGAGAAAGATAAAATTCGATTATAT	948
QY	316	GlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGly	335
Db	949	GGATCCCAAGCAGGCCCTTCTTGGTTATCAAAACCTGTCACCTGGACAGAGTCTTATTGGC	1008
QY	336	SerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValProLeuMet	355
Db	1009	-----CTTGGCTCACCATGGAGCATCATCAACCAAGCATGCCCTCATG	1056
QY	356	AspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly-----	373
Db	1057	GATCCTCTGGTGACACTGTTGGTAGCATTCATGAGAAGCTCCCCGAGACAGGACGAAGA	1116
QY	374	GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln	393
Db	1117	GGAAGCATGCGAAGCACTCTGTTTCCAAATTTTGGAAAGCATGTTTCAGCACTGCTGAGCCG	1176
QY	394	HisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAla	413
Db	1177	CATGCTAAATTTGAACAATGGGATGAAGAAAGCTTACAAAGGGAACGTGAGGACTACATG	1236
QY	414	SerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGln	433
Db	1237	TCAGATGCAACCCGTGGGACTCCGATGATAATTTGCACAGTCCCTTAATCTCAGCCAA	1296
QY	434	AlaThrGlyAlaGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeu---	452
Db	1297	ACAACAAGCCTTGAA--AAAGACTTACCTCCTCCTCCTTCCCATGGCAGTATCCTTGGC	1353
QY	453	SerMetArgArgGlnThrLeuLeuGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly	471
Db	1354	AGCATGAGCGCTCAGTAGTCTCATGCAAGGTCAGGTGAGCAAGGTGGTAGTACAGGT	1413
QY	472	IleGlyGlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArg	491
Db	1414	ATTGGTGGTGGCTGGCAACTGGCATGGAATGGACTGATAAA--GGTGAGGATGGAAA	1470
QY	492	LysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArg	511
Db	1471	CAACAAGGAGGGTTAAAGGATTTATTTATCATGAGGAGGGAGTTTCTGCATCTCGTCGT	1530
QY	512	GlySerIleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHis	531
Db	1531	GGATCCATTGTATCGATTCCCGGTGAAGGC-----GAATTTGTCCAG	1572
QY	532	AlaAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMet	551
Db	1573	GCTGTGCTGCTTGGTAAGCCAAACCCGCTCTTTACTCCAAGGAGCTTATTGATGGACACCCA	1632
QY	552	SerAspAlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAsp	571
Db	1633	GTTGGGCTGCAATGGTTTCAACCATCTGAGACAGCTTCAAAAGGGGCCAAGTTGGAAAGCT	1692
QY	572	LeuPheGluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGln	591
Db	1693	CTTCTTGAACCAAGGGTTAAGCATGTCATTGGTTGTTGGAGTTGGAATACAAATACTTCAG	1752
QY	592	GlnPheAlaGlyIleAsnGlyValLeuTyrThrProGlnIleLeuGluGlnAlaGly	611
Db	1753	CAGTTTTTCAGGATAAATGGGGTTCTATATATACACACCTCAATCCTTGAAGAGGCCGGT	1812

QY 612 ValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSer 631
 Db 1813 GTTGAAGTTCTTCTTTCAGATATAGGCATTGGCTCAGAGTCGGCATCATCTTATCAGT 1872
 QY 632 SerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeu 651
 Db 1873 GCTTTCACAACTTCTTGATGCTTCCCTGTATAGCGGTAGCCATGAGCTCATGGATGT 1932
 QY 652 SerGlyArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIle 671
 Db 1933 TCAGGCAGAAGGCAGTTGCTACTTACTACAAATCCCGCTGCTGATTGTGCTCACTCATATT 1992
 QY 672 LeuValValSerAsnLeuLeuAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrVal 691
 Db 1993 TTGGTCATTGGAAAGCCCTGGTAAATTTTGGCAATGTCGCCCATGCAATCTCAACAGTA 2052
 QY 692 SerValIleValThrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeu 711
 Db 2053 TCGCTTGTTGTTTATTTCTGCTGCTTTGTGATGGGTTATGSAACCAATCCAAACATCCTT 2112
 QY 712 CysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThr 731
 Db 2113 TGCTCAGAGATTTTCCCACTAGGTGCGGCTCTGCATTGCTATCTGTGCAATTAGTG 2172
 QY 732 PheTrpIleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGly 751
 Db 2173 TTCTGGATTGGAGACATCATCATCACTACTCGCTGCTGTGATGCTCGGCTCTTTAGGA 2232
 QY 752 LeuAlaGlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPhe 771
 Db 2233 CTGGTGGTGTATTGCGCATTTACGCAGTTGTTGTTTTCATCTCGTGGATATTGTGTTT 2292
 QY 772 LeuLysValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaVal 791
 Db 2293 TTGAAGGTTCCAGAAACAAAGGGCATGCCCTTGAAGTCATCTCTGAATTTCTTCTGTT 2352
 QY 792 GlyAlaLysGlnAlaAlaAla 798
 Db 2353 GGAGCAAAGCAGGCTGCTTCT 2373

RESULT 8

US-10-051-909-7
 ; Sequence 7, Application US/10051909
 ; Publication No. US20020199217A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Allen, Steve
 ; APPLICANT: Helentjaris, Tim
 ; APPLICANT: Hitz, Bill
 ; APPLICANT: Kinney, Tony
 ; APPLICANT: Tingey, Scott
 ; TITLE OF INVENTION: Plant Sugar Transport Proteins
 ; FILE REFERENCE: BB1163 US CIP
 ; CURRENT APPLICATION NUMBER: US/10/051,909
 ; CURRENT FILING DATE: 2002-01-17
 ; PRIOR APPLICATION NUMBER: 60/083,044
 ; PRIOR FILING DATE: April 24, 1998
 ; NUMBER OF SEQ ID NOS: 38
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 7
 ; LENGTH: 2601
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 US-10-051-909-7

Alignment Scores:
 Pred. No.: 7.15e-255 Length: 2601
 Score: 2674.00 Matches: 522
 Percent Similarity: 82.20% Conservative: 92
 Best Local Similarity: 69.88% Mismatches: 115
 Query Match: 65.62% Indels: 18
 DB: 13 Gaps: 8

US-10-051-909-32 (1-800) x US-10-051-909-7 (1-2601)
 QY 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrp 75
 Db 175 ATGAAAGGTGCGCTCTTGTGTTGTTATGCGCTTCCATTTGGTAATTTCTTCCCAAGGATGG 234
 QY 76 AspAsnAlaThrIleAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsn 95
 Db 235 GATAATGCTACCATCGCCGGGGCTAATGTTTACATTAAGAAAGACCTTGTCTTTGGGAACA 294
 QY 96 GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115
 Db 295 -----ACTATGAAAGGCTTGTGGTGGCATGTCCCTGATTGGAGCAACGGTAATCACC 348
 QY 116 ThrPheSerGlyProLeuSerSerSerIleGlyArgArgProMetLeuIleLeuSerSer 135
 Db 349 ACATGCTCTGCTCTATAGCGGATGGCTCGGTGCGGCGCCCATGATGATATCTCATCT 408
 QY 136 IleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyrValLeuLeu 155
 Db 409 GTGCTCTATTCTTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 468
 QY 156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr 175
 Db 469 TTGGCGAGGCTACTTGTATGGATTGGGATTGGCTTGGCTTGGTGGTGGTGGTGGTGGTGGTGG 528
 QY 176 IleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSer 195
 Db 529 ATATCTGAACCGCGCGCTGAAATAAGGGGGTCTGTGAATACGCTTCTCTCAGTTTCTCAGT 588
 QY 196 GlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSer 215
 Db 589 GGCTCTGGAGGAATGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 648
 QY 216 ProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeu 235
 Db 649 CCTAGCTGGAGGCTCATGCTTGGGTTCTGTCTATTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 708
 QY 236 ThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGlu 255
 Db 709 ACCATTTTCTTGGCGAGTCTCTCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 768
 QY 256 AlaLysLysValLeuGlnLysLeuArgGlyLysAspAspValSerGlyGluLeuSerLeu 275
 Db 769 GCTAAGAAGGTGCTCCAAAGATGCGCGGAAGGAGGATGTGTGAGGCGAGATGGCATTG 828
 QY 276 LeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIleGlyPro 295
 Db 829 CTGGTTGAAGGTCTCGGGATTGGGGGTGATACATCTATCGAAGAGTACATAATTGGCCCT 888
 QY 296 AlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyr 315
 Db 889 GCTGACGATGTGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 948
 QY 316 GlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGly 335
 Db 949 GGATCCCAAGCAGGCTTCTTGTGTTATCAAAACCTGTCACTGGACAGAGTCTATTGGC 1008
 QY 336 SerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValProLeuMet 355
 Db 1009 -----CTTGGCTCACACCATGGAAGCATCATCAACCAAGCATGCCCCCTCATG 1056
 QY 356 AspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly----- 373
 Db 1057 GATCTCTGTTGACACTGTTTGGTAGCATTCATGAGAAAGCTCCCCGAGACAGAGCAAGA 1116
 QY 374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393
 Db 1117 GGAAGCATGCGAAGCACTCTGTTTCCAAATTTTGGAAAGCATGTTTCAGCACTGTGAGCCG 1176
 QY 394 HisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAla 413
 Db 1177 CATGCTAAATTTGAACAATGGATGGATGAAGAAAGCTTACAAAGGGAACGCTGAGGACTACATG 1236

QY 196 GlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSer 215
DB 421 GGCTCTGGAGGGATGTTCTTATCTTACTGTATGTTTTCGGAATGTCGTGATGCCATCA 480
QY 216 ProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeu 235
DB 481 CCTAGCTGGAGATTGATGCTTGGTGTCTTTTTCATCCCTTCCCTTGTCTTTTCTTCTC 540
QY 236 ThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGlu 255
DB 541 ACGGTCCTCTTCTTCCCGAGTCCCGCAAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 600
QY 256 AlaLysLysValLeuGlnLysLeuArgGlyLysAspValSerGlyGluLeuSerLeu 275
DB 601 GCAAAGCGGGTCTTTCAGAGCTGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 660
QY 276 LeuLeuGluGlyLeuGluValGlyAspThrSerIleGluGluTyrIleIleGlyPro 295
DB 661 TTGGTTGAGGGTCTTGAATTGGAGGTGAACAACCATAGAGGATATATAATTGGTCCC 720
QY 296 AlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyr 315
DB 721 GCGGATGAAGTTACTGATGATCATGATATAGTGTGGTAAAGTCAAAATTAAGTTATAT 780
QY 316 GlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGly 335
DB 781 GGTGCAAGAAGAGGGCTGAGTGGTGTGCTAGGCCAGTCAAAGGA-----GGA 828
QY 336 SerValLeuSerLeuAlaSerArgHisGlySer---MetValAsnGlnSerValProLeu 354
DB 829 AGCACTATGAGTGTGTTGCTCGCCATGGAAGTACAATGAGCAGGAGGAGGCTCATTTG 888
QY 355 MetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGlyGly 374
DB 889 ATTGATCTCTTGTACACTGTTTGGGAGCGTTACGAGAGAGTATGCGGACACT---GGA 945
QY 375 SerMetArgSerThrLeuPheProAsnGlySerMetPheSerValThrAspGlnHis 394
DB 946 AGCATGAGGAGTGCCTTGTTCACACATTTTGGGAGTATGTTTTCAGTGTGGAGGAATCAA 1005
QY 395 AlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAlaSer 414
DB 1006 CCAGACATGAAGATTGGGATGAAGAGATCTTGTGGAGAGGAGGAGGATATCCATCC 1065
QY 415 AspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAla 434
DB 1066 GAC---CATGGAGATGATTCTGAAGATGATCTTCTATCTCCGTTGATCTCACGTCACACG 1122
QY 435 ThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeuSerMet 454
DB 1123 ACAAGCATGGAG---AAAGACATGCTCCTCACACTGCTCATGGAACCTCTTCTACCTTCAGA 1179
QY 455 ArgArgGlnThrLeuLeuGlyGluGlyAspGlyValSerSerThrAspIleGlyGly 474
DB 1180 CATGGAAGTCAAGTGCAGGAGCTCAAGGGGAAGGAGCGGTAGTATGCGGATGGAGGT 1239
QY 475 GlyTrpGlnLeuAlaTrpLysTrpSerGluLysGlyGluAsnGlyArgLysGlyGly 494
DB 1240 GGATGCAAGTGGCATGGAATGGACGGAAGAGAGATGAATCGGGACAGAAAGAA--- 1296
QY 495 GlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArgGlySerIle 514
DB 1297 -----GAAGGTTTCCAGGATCTCGACGTGGCTCAATT 1329
QY 515 ValSerLeuProGlyGlyAspValPheGluGlySerGluPheValHisAlaAlaAla 534
DB 1330 GTTTCATTGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1386
QY 535 LeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAspAla 554
DB 1387 TTGGTTAGCCCAACCGCTCTTTATTTCCAAAGACCTTCTCAAAGAACATACAAATTGGTCC 1446
QY 555 AlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuPheGlu 574

DB 1447 GCTATGGTACATCCATCCGAA---ACAACTAAAGGTCAATTGGCATGATCTTCATGAT 1503
QY 575 ProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnPheAla 594
DB 1504 CCTGGAGTCAAGCGTGCATTAGTCGTAGAGTTGGACTTCAAATACTTCAGCAGTTCTCA 1563
QY 595 GlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyValAlaVal 614
DB 1564 GGCATCAACCGGAGTCTTTACTACACACCGCAATCCTTGAGCAGCGGGTGTGGGATC 1623
QY 615 IleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThr 634
DB 1624 CTACTATCGAACATGGGATTTAGTCTCTCTCCTCAGCATCCCTTACTTATAAGTGCATTGACA 1683
QY 635 ThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArg 654
DB 1684 ACCTTTGTGATGTACCTGCAATAGCTGTGCAATGAGGCTCATGGATCTTTCTGTGTCGA 1743
QY 655 ArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuValVal 674
DB 1744 AGGACCTTGCCTTACCACGATACCAATCCTGATAGCATCTTATTTGTTTAGTAATC 1803
QY 675 SerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIle 694
DB 1804 TCAAAATCTTGTTCACATGAACAGCATGTGTCACGCGGTCTTATCAACCGTAAGCGTTGTG 1863
QY 695 ValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAlaGlu 714
DB 1864 CTCTACTTCTGCTTCTCGTGATGGGTTTGGTCTCTGCTCCAAACATCCTCTGTTCAGAG 1923
QY 715 IlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIle 734
DB 1924 ATTTTCCAACTCGAGTCCGCGGAATCTGCATCGCCATCTGGCAGCTCACCTTCTGGATC 1983
QY 735 GlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGly 754
DB 1984 TGTGACATATCGTCACTTACAGTCTCCCCGCTGCTGCTCAAATCCATTTGAGCTGGT 2043
QY 755 ValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeuLysVal 774
DB 2044 GTGTTTGAATGTACGCAATCGTATGTGTCATTTGCTGGGCTTTGTGTTTCAATAAGTC 2103
QY 775 ProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAlaLys 794
DB 2104 CCGGAACTAAAGCATGCCACTTGAAGTCATCAGAGTTCTTTTCTGTTGGAGCTAGA 2163
QY 795 GlnAlaAlaAla 798
DB 2164 CAAGCTGAAGCT 2175

RESULT 10

US-09-938-842A-1315

; Sequence 1315, Application US/09938842A

; Publication No. US20040009476A9

; GENERAL INFORMATION:

; APPLICANT: Harper, Jeff

; APPLICANT: Kreps, Joel

; APPLICANT: Wang, Xun

; APPLICANT: Zhu, Tong

; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING

; TITLE OF INVENTION: SAME, AND METHODS OF USE

; FILE REFERENCE: SCRIPI300-3

; CURRENT APPLICATION NUMBER: US/09/938,842A

; CURRENT FILING DATE: 2001-08-24

; PRIOR APPLICATION NUMBER: US 60/227,866

; PRIOR FILING DATE: 2000-08-24

; PRIOR APPLICATION NUMBER: US 60/264,647

; PRIOR FILING DATE: 2001-01-16

; PRIOR APPLICATION NUMBER: US 60/300,111

; PRIOR FILING DATE: 2001-06-22

; NUMBER OF SEQ ID NOS: 5379

; SEQ ID NO 1315


```
QY 498 gValTyrLeuHisGlnGluValProGlySerArgArgGlySerIleValSerLeuPr 518
Db 1579 GATTATTATCATGAGGAGGAGTTTCTGTCATCTCGTGGATCCATTGTATCGATTCC 1638
QY 518 oGlyGlyAspValPheGluGlySerGluPheValHisAlaAlaAlaLeuValSerG1 538
Db 1639 CGGTGAAGGC-----GAATTTGTCAGGCTGCTGCCCTTGGTAAGCCA 1680
QY 538 nSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHi 558
Db 1681 ACCCGCTCTTTACTCCAAGGAGGCTTATTGATGGACACCCAGTTGGCCCTGCAATGTTCA 1740
QY 558 sProSerGluValAlaAlaLysGlySerArgTyrPheLeuPheGluProGlyValAr 578
Db 1741 CCCATCTGAGACAGCTTCAAAGGGCCCAAGTTGGAAGCTCTTCTTGAACCCAGGGGTTAA 1800
QY 578 gArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnPheAlaGlyIleAsnG1 598
Db 1801 GCATGCAATGGTTGTTGGAGTTGGAATACAAATACCTTCAGCAGTTTTCAGGGATAAATGG 1860
QY 598 yValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyValAlaValIleLeuSerLy 618
Db 1861 GGTTCTATATTACACACCTCAATCCTTGAAGAGGCGGTGTTGAAGTTCTTCTTTCAGA 1920
QY 618 sPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMe 638
Db 1921 TATAGGCATTTGGCTCAGAGTCGGCATCATCTCTTATCAGTGTCTTTCACAACTTCTTGAT 1980
QY 638 tLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeuLe 658
Db 1981 GCTTCCCTGTATAGCGGTAGCCATGAAGCTCATGGATGTTTCAGGCAGAGGAGTGTGT 2040
QY 658 uLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuValValSerAsnLeuI1 678
Db 2041 ACTTACTACAATCCCGCTGTGATGGTCACTCAATATTGTTGTCATTGGAAGCCTGGT 2100
QY 678 eAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyrPheCy 698
Db 2101 AAATTTGGCAATGTGCGCCCATGCAGCAATCTCAACAGTATGCGTTGTGGTTATTCTG 2160
QY 698 sCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePheProTh 718
Db 2161 CTGCTTTGTGATGGGTTATGGACCAATTCCAAACATCTCTTGTCTCAGAGATTTCCTCC 2220
QY 718 rArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIleGlyAspIleI1 738
Db 2221 TAGGGTGGTGGCCCTCTGCAATGCTATCTGTGCAATAGTGTCTGGATTGGAGACATCAT 2280
QY 738 eValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSerI1 758
Db 2281 CATCATATACTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2340
QY 758 eTyrAlaValValCysLeuIleSerPheValPheLeuLeuLysValProGluThrLy 778
Db 2341 TTACGAGTTGTTTGTTCATCTCGTGGATATTGTTGTTTGAAGGTTCCAGAAACAAA 2400
QY 778 sGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAlaLysGlnAlaAla1 798
Db 2401 GGGCATGCCCTTGAAGTCATCTCTGAATCTCTTCTGTTGGAGCAAGCAGGCTGCTTC 2460
QY 798 a 798
Db 2461 T 2461
```

RESULT 12

US-10-425-114-7958
; Sequence 7958, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E

; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 7958
; LENGTH: 2405
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700728773_FLI
US-10-425-114-7958

Alignment Scores:
Pred. No.: 6.49e-242 Length: 2405
Score: 2543.00 Matches: 508
Percent Similarity: 79.54% Conservative: 83
Best local Similarity: 68.37% Mismatches: 120
Query Match: 62.40% Indels: 32
DB: 16 Gaps: 10

US-10-051-909-32 (1-800) x US-10-425-114-7958 (1-2405)

```
QY 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrp 75
Db 76 ATGAGCGGAGCTGTTCTTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 135
QY 76 AspAsnAlaThrIleAlaAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsn 95
Db 136 GATAATGCCACCATTCAGGATCTATTTTGTACATAAAGAGGGAGTTTAAATTACAAAGT 195
QY 96 GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115
Db 196 GAACCCACAATAGAGGTCTAATTGTGGCCATGTCTATTTGGAGCCACTGCTGTTACG 255
QY 116 ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer 135
Db 256 ACATGCTCTGGACCCCTTATCAGACTTGTAGCCCGGCTCTATGTTGATCATCTCTCT 315
QY 136 IleLeuTyrPhePheSerGlyLeuIleMetLeuTyrSerProAsnValTyrValLeuLeu 155
Db 316 ATCTTTTATTTTGTGTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 375
QY 156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr 175
Db 376 TTTGCAAGGCTTTTAGATGAGATTAGGCAATTGGTTTGGCGGTCACTTGGTACCTCTTTAC 435
QY 176 IleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSer 195
Db 436 ATATCCGAGACAGCTCCACCTGAGATTAGGGGATTACTAAATACTCTTCCGCACTTCACT 495
QY 196 GlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSer 215
Db 496 GGTCTGCTGGAATGTTCTTTCTCTACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 555
QY 216 ProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeu 235
Db 556 CCAAACTGGAGACTCATGTAGGTGTTCTTTCAATTCCTCTCTCATTTTATTTGCACACT 615
QY 236 ThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGlu 255
Db 616 ACACTATTCTTCTTGCCTGAATCTCCAAGATGGTGTGTCAGCAAAAGCGGATGCTAGAG 675
QY 256 AlaLysLysValLeuGlnLysLeuArgGlyLysAspAspValSerGlyGluLeuSerLeu 275
Db 676 GCCAAGAAGGTTTTTGCACGACTTCGCGGCAGACAAGATGTGCTGGTGAGATGGCTTTA 735
QY 276 LeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIleGlyPro 295
Db 736 CTAGTTGAGGGTCTTGGAGTTGGGGTGTATACCGTTATAGAGAGTACATAAATTAGTCCA 795
```


Qy 296 AlaThrGluAlaAlaAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyr 315
Db 796 GCCAATGAATTCAGT-----GATCAGAGGAGATCAAATTGTAT 834
Qy 316 GlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGly 335
Db 835 GGGACAGCAGAGGTCAATCCTGGATTGCCAAACCTGTCACTGGCAAAAGTTCCATTGGC 894
Qy 336 SerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValProLeuMet 355
Db 895 -----CTTGATCTCGGAAGGGAAGCATGGCAAAATCAAAGTGCT---CTAGTG 939
Qy 356 AspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGlyGlySer 375
Db 940 GACCCCTCTAGTAAGCTCTTTGGTAGTGTCCATGAGAAGCTCCCAAGAACAGGA----- 993
Qy 376 MetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGlnHisAla 395
Db 994 -----AGCACACTTTTCCACACTTTTGGAAAGCATGTTTAGTGTGGGGGAAATCAACCT 1047
Qy 396 LysAsnGluGlnTrpAspGluAsnLeuHisArgAspAspGluGluTyrAlaSerAsp 415
Db 1048 AGGAATGAAGATTGGGATGAAGAAAGCATTCGCCAGAGAGGGAGATGATTATGTATCTGAT 1107
Qy 416 GlyAlaGly---GlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAla 434
Db 1108 GCTGCTGACACCGACGATTCTGATGACAATTTGCAGAGTCCATTCATGATCAGCA 1167
Qy 435 ThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeuSerMet 454
Db 1168 ACAAGTCCGGAA---AGGGACATGCTCTGCTGCCAGGGA-----AGCATG 1212
Qy 455 ArgArgGlnThrLeuLeuGlyGluGlyGlyAspGlyValSerSerThrAspIleGlyGly 474
Db 1213 AGGCAAGGTAGTCTTTTG-----CAAGGAGAAACCTGCTGGAAACTCAGGGATTGGTGGT 1266
Qy 475 GlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArgLysGluGly 494
Db 1267 GGATGGCAGCTAGCATGGAATGGTCTGTAACAGAGGTGTT----- 1308
Qy 495 GlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArgGlySerIle 514
Db 1309 ---TTCAAGAGATATATTTGCACCACAGAGGTGCTCTGGATCTAGACGTGGTCCATC 1365
Qy 515 ValSerLeuProGly---GlyGlyAspValPheGluGlySerGluPheValHisAlaAla 533
Db 1366 ATTTCACTTCCAGGCTGTGATGCGCCGCACACTAACAGATGGTGAGATTGTCCAGGCTGCC 1425
Qy 534 AlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAsp 553
Db 1426 GCTCTAGTGAGTCAATCAGCCCTTTACACAAGGAGCTTATGTCATCAGCAGCCAGTTGGA 1485
Qy 554 AlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuPhe 573
Db 1486 CCAGCCATGATTATCATCCATCCCAACAGCTGCAAAAGGGCCGAGTTGGAGTGACCTTTTA 1545
Qy 574 GluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnGlnPhe 593
Db 1546 GAGCCTGGAGTGAAGCATGCATGATTGTGGGGGTGGGATTCAAATTTCTTCAACAGTTC 1605
Qy 594 AlaGlyIleAsnGlyValLeuTyrTrpThrProGlnIleLeuGluGlnAlaGlyValAla 613
Db 1606 TCTGGTATAAATGGGGTGTCTCTACTATACACCTCAAATTTCTTGACAGGCAGGTGTGGT 1665
Qy 614 ValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeu 633
Db 1666 TATCTTCTTTCAAACCTGGGCCCTCGGTTCTACTTCTGCGTCTTTTCTTATCAGCAGCGTG 1725
Qy 634 ThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGly 653
Db 1726 ACAACCTTGTGATGTTACCCCTGTATAGCTGTGGCCATGAGGCTCATGGATATATCTGGC 1785

Qy 654 ArgArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuVal 673
Db 1786 AGAAGACCCCTTGCTGCTACTACATCCCGCTTGATAGTATCACTTCTCATATTAGTA 1845
Qy 674 ValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerVal 693
Db 1846 ATAGGGAGTCTTGGAATGGACAGTACTATAAATGCTTTTATCTCAACTTCGAGTGT 1905
Qy 694 IleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAla 713
Db 1906 ATTGTCTACTTCTGTGTTTTCGTATGGGCTTTGGGCCAATTCTCTAATAATCTTGTTC 1965
Qy 714 GluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrp 733
Db 1966 GAGATCTTCCCTACTCGAGTTCGTGGTCTATGATAGTATATATGTGCTTCTTACATTTGG 2025
Qy 734 IleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAla 753
Db 2026 ATATGTGACATCATAGTACCTACTCGCTCCAGTTATGCTTAATTCGTGGGCTTGGT 2085
Qy 754 GlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeuLys 773
Db 2086 GGTGCTTTGGTATGTATCGGTTGTGTGCATCATAGCTTGGGTGTTGTCTTCTTAAAA 2145
Qy 774 ValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAla 793
Db 2146 GTTCCAGAAACCAAGGCGATGCCACTGGAAGTATCATTTGAGTTCTTCTCTGTCGGG 2205
Qy 794 LysGlnAla 796
Db 2206 AAGCAAGCT 2214
RESULT 13
US-10-424-599-108435
; Sequence 108435, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 108435
; LENGTH: 2573
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_68931C.1
US-10-424-599-108435

Alignment Scores:
Pred. No.: 5,86e-236 Length: 2573
Score: 2483.50 Matches: 505
Percent Similarity: 76.69% Conservative: 84
Best Local Similarity: 65.76% Mismatches: 121
Query Match: 60.94% Indels: 58
DB: 16 Gaps: 12

US-10-051-909-32 (1-800) x US-10-424-599-108435 (1-2573)

Qy 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrp 75
Db 37 ATGAGCGGAGCTGTTGTTGCTGTGGTGTGCTATCGGCAACTTGTGCTCAAGGATGG 96
Qy 76 AspAsnAlaThrIle----- 80
Db 97 GATAATGCCACCATTCAGGTTCATATTTTATTCGGTTAGCCAATTAACATAACATTCAA 156

QY 81 -----AlaAlaValLeuTyrIleLysLysGluPhe 91
Db 157 CCAAGTCTAATTGGCTTTCATGCTATGACGAGTCTATTTGTACATAAAGAGGAGTTT 216
QY 92 GlnLeuGlnAsnGluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAla 111
Db 217 AATTACAAAGTGAACCCACAATAGAAGGTCTAATTGTGGCCATGTCACTTATTGGAGCC 276
QY 112 ThrIleValThrThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeu 131
Db 277 ACTGTGTTACGACATGCTCTGGACCTTATCAGACTTGCTAGGCCGGCGTCTATGTTG 336
QY 132 IleLeuSerSerIleLeuTyrPhePheSerGlyLeuIleMetLeuTyrPheSerProAsnVal 151
Db 337 ATCATCTCCTCTATCCTTTATTTGTTGGTCTCTGGTCATGCTATGGTCTCCAAATGTT 396
QY 152 TyrValLeuLeuLeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeu 171
Db 397 TATATTCTCCTCTTTGCAAGGCTTTTAGATGGATTAGGCATTGGTTTGGCGGTACCTTG 456
QY 172 ValProLeuTyrIleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeu 191
Db 457 GTACCTCTTTACATATCCGACACAGCTCCACCTGAGATTAGGGATTACTAAATACTCTT 516
QY 192 ProGlnPheSerGlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSer 211
Db 517 CCGCAGTTCACTGGTTCTGCTGGAATGTTCTTTTCTACTGTATGGTCTTTGCTATGTCA 576
QY 212 LeuSerProSerProAspTyrArgIleMetLeuGlyValLeuAlaIleProSerLeuPhe 231
Db 577 CTCACCAAGGCCCAAAATTGGAGACTCATGTTGGGTGTTCTTTCAATTCCCTCTCTCAT 636
QY 232 PhePheGlyLeuThrIlePheTyrLeuProGluSerProArgTyrLeuValSerLysGly 251
Db 637 TATTTTGCACTCACGCTATTCTTCTGCTGGAATGTTCTTTTCTACTGTATGGTCTTTGCTATGTCA 576
QY 252 ArgMetAlaGluAlaLysLysValLeuGlnLysLeuArgGlyLysAspValSerGly 271
Db 694 CGGATGTAGAGGCCAAGAAGGTTTGCACACGACTTCGCGGCAGACAAGATGTCGCTGT 753
QY 272 GluLeuSerLeuLeuGluGlyLeuGluValGlyAspThrSerIleGluGluTyr 291
Db 754 GAGATGCTTACTAGTTGAGGTCTTGGAGTTGGGCTGATACCGCTATAGAAGAGTAC 813
QY 292 IleIleGlyProAla-ThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluG1 311
Db 814 ATAATTGGTCCAGCTCAATGAATTCAGT-----GAGCAGAGCA 852
QY 311 nIleThrLeuTyrGlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyPr 331
Db 853 GATCAATTTGATGGGACAGAGAGGTGTATCCTGGATTGCTAAACCTGTCACTGGACA 912
QY 331 oIleMetLeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSe 351
Db 913 AAGTTCATTTGGT-----CTTGTATCTCGGAAGGGAAGCATGGCAATCAAAG 960
QY 351 rValProLeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProG1 371
Db 961 T---CGTCTAGTGGACCTCTAGTGAAACTCTTTTGGTAGTGTCCATGAGAACCTCCCAGA 1017
QY 371 nAlaGlyGlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValTh 391
Db 1018 AACAGGA-----AGCACACTTTTCCACACTTTTGGAAAGCATGTTAGTGTGG 1065
QY 391 rAspGlnHisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluG1 411
Db 1066 GGGGAATCAACCTTAGGAATGAAGATTGGGATGAAGAAGCATTTGCCAGAGAGGAGATGA 1125
QY 411 uTyrAlaSerAspGlyAlaGly---GlyAspTyrGluAspAsnLeuHisSerProLeuLe 430
Db 1126 TTATGTATCTGTGCTGTTGACACCGACCGATTTCTGTATGACAATTTGCAGAGTCCATTGAT 1185
QY 430 uSerArgGlnAlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySe 450

Db 1186 CTACAGTCAAGCAACAAGTGGGAA---AGGGACATGCCTGCTCTGCCCCAGGA----- 1237
QY 450 rAlaLeuSerMetArgArgGlnThrLeuLeuGlyGluGlyGlyAspGlyValSerSerTh 470
Db 1238 -----AGCATGAGGCAAGGTAGTCTTTTG-----CAAGGAGAACCCTGCTGGAACCTC 1284
QY 470 rAspIleGlyGlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnG1 490
Db 1285 AGGATTTGGTGGATGGCAGCTAGCATGGAAATGGTCTGAAACAGAGAGGTGTT----- 1339
QY 490 yArgLysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerAr 510
Db 1340 -----TTCAAGAGAAATATATTTTGACCAAGAAGGTGGTCTCTGGATCTAG 1383
QY 510 gArgGlySerIleValSerLeuProGly---GlyGlyAspValPheGluGlySerGluPh 529
Db 1384 ACGTGGTCCATCATTTTCAGGCTGTGATGCGCCGACACTAACAGATGGTGAGAT 1443
QY 529 eValHisAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluPr 549
Db 1444 TGTCAGGCTGCCGCTCTAGTGAATCAATCAGCCCTTTACAAACAGGAGCTTATGCATCA 1503
QY 549 oArgMetSerAspAlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTr 569
Db 1504 GCAGCCAGTTGGACCGCATGATTCCATCCCAACAGCTGCAAAAGGCCGAGTTG 1563
QY 569 pLysAspLeuPheGluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnI1 589
Db 1564 GAGTGACCTTTTAGAGCCTGGAGTGAAGCATGCACTGATGTTGGGGTGGGATTCAAAT 1623
QY 589 eLeuGlnGlnPheAlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluG1 609
Db 1624 TCITCAACAGTTCTCTGTATATAAATGGGTGCTCTACTATACACCTCAAATTTCTTGAGCA 1683
QY 609 nAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLe 629
Db 1684 GGCAGGTGTTGGTTATCTTCTTTCAAACCTGGGCTCGGTCTCTACTTCTGCGTCTTTCT 1743
QY 629 uIleSerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMe 649
Db 1744 TATCAGCAGCGTGACAACCTTGTGTATGTTACCCCTGTATAGCTGTGGCCATGAGGCTCAT 1803
QY 649 tAspLeuSerGlyArgArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLe 669
Db 1804 GGATATATCTGGCAGAGAAGACCTTGTCTGTACTACATAAATCCCGTCTTGATAGTATCACT 1863
QY 669 uValIleLeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSe 689
Db 1864 TCITCATATTAGTAATAGGAGTCTTTGTGGAATTGGACAGTACTATAAATGCTTTTATCTC 1923
QY 689 rThrValSerValIleValTyrPheCysCysPheValMetGlyPheGlyProIleProAs 709
Db 1924 AACTTCGAGTGTATTGTTCTACTTCTGTTGTTTCGTCATGGGCTTTGGGCCAATTCCTAA 1983
QY 709 nIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAl 729
Db 1984 TATCCTTTGTTTCAGAGATCTTCCCTACTCGAGTTCGTTGCTGTATGATATATGTGC 2043
QY 729 aPheThrPheTrpIleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAl 749
Db 2044 TCTTACATTTTGGATATGTGACATCATAGTCACTACTCGTCCCAGTTATGTTAATTC 2103
QY 749 aIleGlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPh 769
Db 2104 TGTGGGCTTGGTGGTGTCTTTGGTATGATGCGGTGTGTGTCATCATAGTGGTGT 2163
QY 769 eValPheLeuLysValProGluThrLysGlyMetProLeuGluValIleThrGluPhePh 789
Db 2164 TGTCTTCTTAAAGTTCCAGAAACCAAGGGCATGCCACTGGAAAGTTATCATTCAGTTCTT 2223
QY 789 eAlaValGlyAlaLysGlnAla 796

Db 2224 CTCTGTCGGGGCAAAGCAAGCT 2245

RESULT 14

US-10-425-114-29146

; Sequence 29146, Application US/10425114

; Publication No. US20040034888A1

; GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong

; APPLICANT: Zhou, Yihua

; APPLICANT: Kovalic, David K.

; APPLICANT: Screen, Steven E

; APPLICANT: Tabaska, Jack E

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53313)B

; CURRENT APPLICATION NUMBER: US/10/425,114

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 73128

; SEQ ID NO 29146

; LENGTH: 2271

; TYPE: DNA

; ORGANISM: Zea mays

; FEATURE:

; OTHER INFORMATION: Clone ID: LIB83-005-C10_FLI

US-10-425-114-29146

Alignment Scores:

Pred. No.:	5.7e-232	Length:	2271
Score:	2442.50	Matches:	472
Percent Similarity:	83.66%	Conservative:	76
Best Local Similarity:	72.06%	Mismatches:	100
Query Match:	59.94%	Indels:	7
DB:	16	Gaps:	6

US-10-051-909-32 (1-800) x US-10-425-114-29146 (1-2271)

QY 143 LeuIleMetLeuTrpSerProAsnValTyrValLeuLeuAlaArgPheValAspGly 162

Db 3 CTGGTGATGCTCTGGCGCCCAAGTGTGTACATCTTGTCTCTCGCAAGGCTCATTGATGGG 62

QY 163 PheGlyIleGlyLeuAlaValThrLeuValProLeuTyrIleSerGluIleAlaProSer 182

Db 63 TTCGGTATCGGTTTGGCGGTACACATGTTCTCTCTACATCTCCGAGACTCGCGCCGACA 122

QY 183 GluIleArgGlyLeuLeuAsnThrLeuProGlnPheSerGlySerGlyGlyMetPheLeu 202

Db 123 GACATTCGTGGTGTGTAACACGTTGCCGCAGTTCAGTGGGTGAGGAGGATGTTCCCTC 182

QY 203 SerTyrCysMetValPheGlyMetSerLeuSerProSerProAspTrpArgIleMetLeu 222

Db 183 TCCTACTGTCATGGTGTGTTGGGATGTCCTCATGCCCAACCTGATTGGAGGCTCATGCTT 242

QY 223 GlyValLeuAlaIleProSerLeuPhePheGlyLeuThrIlePheTyrLeuProGlu 242

Db 243 GGAGTTCTGTCGATCCCGTCACATTATTTACTTTGGACTGACTGCTTCTACTTGCCTGAA 302

QY 243 SerProArgTrpLeuValSerLysGlyArgMetAlaGluAlaLysLysValLeuGlnLys 262

Db 303 TCACCAAGGTGGCTTGTGAGCAAAAGGAGGATGGCTGAGGGCAAGAGAGTGTTCGCAAGG 362

QY 263 LeuArgGlyLysAspAspValSerGlyGluLeuSerLeuLeuGluGlyLeuGluVal 282

Db 363 CTGCGGGGAAGAGAAGATGTCACAGGGGAGATGGCTCTTCTAGTTGAAGGTTTGGGGGTC 422

QY 283 GlyGlyAspTrpSerIleGluGluTyrIleIleGlyProAlaThrGluAlaAlaAspAsp 302

Db 423 GGTAAAGATACACGTATTGAAGAATACATAATTGGTCCCGATGATGAACTTGTGTATGAA 482

QY 303 LeuValThrAspGlyAspLysGluGlnIleThrLeuTyrGlyProGluGluGlnSer 322

Db 483 GGGCTGGCTCCA---GATCCAGAGAAGATCAAACTATATATGACCTTGACCTGAAGAAGGCCTATCT 539

QY 323 TrpIleAlaArgProSerLysGlyProIleMetLeuGlySerValLeuSerLeuAlaSer 342

Db 540 TGGGTTGCCGACCTGTTCCGGGACAAAGTGCTCTTGGAAAGCGCTTAGGTCTCATCTCT 599

QY 343 ArgHisGlySerMet--ValAsnGlnSerValProLeuMetAspProIleValThrLeu 361

Db 600 CGTCATGGGAGTATGGCGGCTAGTCAGGGTAAGCCCTCGTGGATCCTATGGTCACTCT 659

QY 362 PheGlySerValHisGluAsnMetProGlnAlaGlyGlySerMetArgSerThrLeuPhe 381

Db 660 TTCGGAAGTGTTCATGAAAGATGCTGTAGATCATGGGGAGCATGAGGAGCACATTGTTT 719

QY 382 ProAsnPheGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTrpAsp 401

Db 720 CCCAACTTGGCAGCATGTTTAGTGTTCGCCACCAGCAGCAGGTGAAAGCTGACTGGGAC 779

QY 402 GluGluAsnLeuHisArgAspAspGluGluTyrAlaSerAspGlyAlaGlyGlyAspTyr 421

Db 780 GCCGAGAGT---CAAAGGAAGGTGAAGATTATGCTTCGGATCATGGTGGCGATGACATT 836

QY 422 GluAspAsnLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAsp 441

Db 837 GAGGATAACCTCCAAAGCCCACCTTATTCTCGTCAGGCAACAAGTGTGAAGGAAGGAG 896

QY 442 IleValHisHisGlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeuGly 461

Db 897 ATCGCT-----GCACCTCATGTTAGCATATTGGGTGCTGTGGGAAGGAGCAGTAGCTTG 950

QY 462 GluGlyGlyAspGlyValSerSerThrAspIleGlyGlyGlyTrpGlnLeuAlaTrpLys 481

Db 951 CAGGGTGGGAGGCAGTAAGCAGCATGGGCATTTGGCGAGGATGGCAGTTGGCGTGGAAA 1010

QY 482 TrpSerGluLysGluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyrLeu 501

Db 1011 TGGACCGAGAGAGAGGGCGAAGATGGGCAAAAGGAAGGTGGCTTCCAGCGTATTTACTTG 1070

QY 502 HisGlnGluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGlyGly 521

Db 1071 CATGAGGAGGGCGTACAAGGCAAC---AGGGGTTCTATATTGTCTATTACCA---GGCGGG 1124

QY 522 AspValPheGluGlySerGluPheValHisAlaAlaLeuValSerGlnSerAlaLeu 541

Db 1125 GATGTTCTCTCGTGGTGAGTTTCATCCAGGCTGCAGCTCTTGTGAGCCAAACAGCTCTT 1184

QY 542 PheSerLysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisProSerGlu 561

Db 1185 TACTCTAAGGAACCTGCTGGAGCAACGTCGTGCTGCTGCTGCGATGATGATCCATCTGAA 1244

QY 562 ValAlaAlaLysGlySerArgTrpLysAspLeuPheGluProGlyValArgAlaLeu 581

Db 1245 GCAGTTACTAAAGTCCAGATGGGCCGACCTATTGTAGCCTGGGGTGAAGCATGCACGTG 1304

QY 582 LeuValGlyValGlyIleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyValLeuTyr 601

Db 1305 TTTGTTGGCATAGGAATACAGATCCTCGCAACAGTTTGTGGCATCAACGGCGTCTCTAC 1364

QY 602 TyrThrProGlnIleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeu 621

Db 1365 TACACTCCTCAAAATCTTGAGCAAGCAGGCGTCGGTGTCTTCTGTGCAACCTCGGCCTT 1424

QY 622 SerSerAlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCys 641

Db 1425 AACGCTTCTTCGGCATCAATCCTCATTAGCGCCCTGACGACCTTACTGATGCTCCCAAGC 1484

QY 642 IleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuGlyThr 661

Db 1485 ATCGGCATTGCGATGAGGCTCATGATATATGTCGGAAGGAGGTTTCTCTCTCTCGCGAGC 1544

QY 662 IleProIleLeuIleAlaSerLeuValIleLeuValValSerAsnLeuIleAspLeuGly 681

Db 1545 ATCCAGTCCCTAATAGTCGGCTACTCGTCCCTGCTGTGTCCCAACATCGTCGACGTGGGG 1604

QY 682 ThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyrPheCysCysPheVal 701


```

Db 1605 GACGTGGCGCACGGCGCTCTCCACGGCCAGCGTCATAGTCTACTTCTGCTTCTCGTC 1664
Qy 702 MetGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePheProThrArgValArg 721
Db 1665 ATGGGTTCGGGCCCGTCCCAACATCTCTCGCGAGAGATCTTCCCAACCAAGGTCGCGC 1724
Qy 722 GlyLeuCysIleAlaIleCysAlaPheThrPheTrpIleGlyAspIleIleValThrTyr 741
Db 1725 GGTCGTGCATCGCCATCTCGGCCCTGGCCCTTCTGGGTGGCGGACATCATCTGACGTAC 1784
Qy 742 SerLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaVal 761
Db 1785 ACTCTCCCGTGATGCTGACGTCGTGGGCTCGCGGCGTCTTTGGGGTGTACGCCGTC 1844
Qy 762 ValCysLeuIleSerPheValPheValPheLeuLysValProGluThrLysGlyMetPro 781
Db 1845 GTGTGCGTCCTAGCCCTCGCGTTCGTGTTTCATCAGGTGCCCGAGACGAGGCGCATGCC 1904
Qy 782 LeuGluValIleThrGluPhePheAlaValGlyAlaLysGlnAla 796
Db 1905 CTCGAGGTTCATCACCGAGTCTTCTCCGTTGGGGCAAGCAAGCC 1949

```

RESULT 15

```

US-09-938-842A-2254
; Sequence 2254, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRI1300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2254
; LENGTH: 2205
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-2254

```

Alignment Scores:

```

Pred. No.: 2.36e-215 Length: 2205
Score: 2275.00 Matches: 462
Percent Similarity: 75.89% Conservative: 111
Best Local Similarity: 61.19% Mismatches: 146
Query Match: 55.83% Indels: 36
DB: 9 Gaps: 14

US-10-051-909-32 (1-800) x US-09-938-842A-2254 (1-2205)

Qy 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrp 75
Db 1 ATGAAGGAGCGACTCTCGTCTCTCGCGCCACAAATCGGCAATTTCTTACAAGGATGG 60
Qy 76 AspAsnAlaThrIleAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsn 95
Db 61 GACAATGCCACCAATGCTGGAGCTATGGTTATATATCAACAAGACTTGAATCTA----- 114
Qy 96 GluProThr---ValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleVal 114
Db 115 ---CCAACCTCTGTTCAAGGTCTTGTGCTGCTATGTCATTGATCGGTGCAACGGTCAFC 171
Qy 115 ThrThrPheSerGlyProLeuSerAspSerIleGlyArgProMetLeuIleLeuSer 134

```

```

Db 172 ACGACTTGCTCAGGACCGATATCTGATTGGCTCGGCAGACGCCCATGCTCATTTATCA 231
Qy 135 SerIleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyrValLeu 154
Db 232 TCAGTTATGATTTCGTCTCGCGGTTTGATAATGTTGGTCTCCCAATGCTATGTTCTG 291
Qy 155 LeuLeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeu 174
Db 292 TGCTTTGCTAGGCTTCTTAATGGGTTTGGTGGCGGCTCGCGTTACACTTGTCCCTGTT 351
Qy 175 TyrIleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPhe 194
Db 352 TACATTTCTGAACCCGCTCTCCGAGATCAGAGGACAGTTAAATACCTCTCCCTCAGTTT 411
Qy 195 SerGlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerPro 214
Db 412 CTGGCTCTGGTGAATGTTTGTCTATCTGATGTTTTCATATGTCCTGAGTGAC 471
Qy 215 SerProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGly 234
Db 472 TCCCTAGCTGGAGAGCCATGCTCGGTGTCCTCTCGATCCCTTCTCTTCTTTT 531
Qy 235 LeuThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMetAla 254
Db 532 CTCACGGTGTGTTTATTTGCCCGAGTCTCTCTGTTGGTGGTTAGTAAAGGAAGATGGAC 591
Qy 255 GluAlaLysLysValLeuGlnLysLeuArgGlyLysAspValSerGlyGluLeuSer 274
Db 592 GAGGCTAAGCGAGTCTTCAACAGTTATGTGGCAGAGAAGATGTTACCGATGAGATGGCT 651
Qy 275 LeuLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGlyTyrIleGly 294
Db 652 TTACTAGTTGAAGGACTAGATATAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 711
Qy 295 ProAlaThrGluAlaAlaAspLeuValThrAspGlyAspLysGlu---GlnIleThr 313
Db 712 TTGGAGGATCATGAAGGTGATGATACATCTTGAACCGTGTGATGAGGATGGACAAATGCGG 771
Qy 314 LeuTyrGlyProGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMet 333
Db 772 CTTTATGGAACCCACGAGAATCAATCGTACCTGTGCTAGACCTGTCCAGAACAA----- 825
Qy 334 LeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValPro 353
Db 826 ---AATAGCTCACCTGGGCTACGCTCTCGCCACGGAAGCTTAGCAAAACCAAGCATGATC 882
Qy 354 LeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly 373
Db 883 CTAAAGATCCGCTCGTCAATCTTTTGGCAGTCTCCACGAGAAGATGCCAGAGCAGGC 942
Qy 374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393
Db 943 GGAAACACTCGGAGTGGGATTTTCCCTCATTTCCGGAAGCATGTTTCTAGTACTACTGCCGAT 1002
Qy 394 -----HisAlaLysAsnGluGlnTrpAsp-----GluGluAsnLeuHisArgAsp 408
Db 1003 GCGCCTCACGGTAAACCGGCTCATTTGGGAAAGGACATAGAGAGCCATTACAACAAGAC 1062
Qy 409 AspGluGluTyrAlaSerAspGlyAlaGlyAspTyrGluAsp-----AsnLeu 425
Db 1063 AATGATGACTATCGGACTGATGATGTTGCGGGTGTATGATGATGATGATGATGATGATG 1122
Qy 426 HisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIleValHisHis 445
Db 1123 CGTAGCCCCCTTAATGTGCGGCCAGACCACAGCATGGAC---AAGGATATGATCCCAT 1179
Qy 446 GlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeuGlyGlyGlyAsp 465
Db 1180 CCTACAAGTGAAGCACTTTAAGCATGAGACGACACAGTACGCTTATGCAA---GGCAAC 1236
Qy 466 GlyValSerSerThrAspIleGlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLys 485

```

Search completed: October 14, 2004, 00:03:19
Job time : 1229.17 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 13, 2004, 12:24:34 ; Search time 18.9209 Seconds
(without alignments)
2138.481 Million cell updates/sec

Title: US-10-051-909-32
Perfect score: 4075
Sequence: 1 IRSGSWLAVQTFFTPDLDRR.....PLEVITEFFAVGAKQAAAKA 800

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 49464 seqs, 25288758 residues

Total number of hits satisfying chosen parameters: 98928

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool_h/US10051909/runat_13102004_123338_19982/app_query.fasta_1.1678
-DB=Pending Patents NA_New -QFMT=fastap -SUFFIX=rnpn -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10051909 @CGN 1 1 271 @runat_13102004_123338_19982 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Pending Patents NA_New.*
1: /cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq.*
2: /cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq.*
3: /cgn2_6/ptodata/1/pna/US07_NEW_COMB.seq.*
4: /cgn2_6/ptodata/1/pna/US08_NEW_COMB.seq.*
5: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq.*
6: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq.*
7: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	2117	52.0	94349	6 US-10-746-294A-96	Sequence 96, Appl
2	354	8.7	2571	7 US-60-613-292-1068	Sequence 1068, Ap
3	352.5	8.7	2573	7 US-60-613-292-899	Sequence 899, App
4	187.5	4.6	582	6 US-10-948-737-10147	Sequence 10147, A
5	181	4.4	1910	7 US-60-613-292-497	Sequence 497, App
6	166	4.1	90149	6 US-10-746-294A-40	Sequence 40, Appl
7	158	3.9	1882	7 US-60-613-292-1274	Sequence 1274, Ap
8	153	3.8	140064	6 US-10-746-294A-47	Sequence 47, Appl
9	142.5	3.5	1629	7 US-60-613-194-105	Sequence 105, App
10	141.5	3.5	411	6 US-10-948-737-5817	Sequence 5817, Ap
11	126	3.1	9888	7 US-60-613-292-399	Sequence 399, App

12	123.5	3.0	473	6 US-10-950-009-135	Sequence 135, App
13	119	2.9	491	6 US-10-950-009-1142	Sequence 1142, Ap
C 14	118	2.9	2016	6 US-10-632-694A-4	Sequence 4, Appli
C 15	118	2.9	2016	6 US-10-632-694A-5	Sequence 5, Appli
C 16	116.5	2.9	31248	6 US-10-485-710-1	Sequence 1, Appli
C 17	116.5	2.9	35359	6 US-10-485-710-2	Sequence 2, Appli
18	112.5	2.8	237	6 US-10-950-009-605	Sequence 605, App
C 19	112	2.7	337	6 US-10-950-009-1043	Sequence 1043, Ap
C 20	112	2.7	1668	6 US-10-411-910B-226	Sequence 226, App
C 21	112	2.7	3621	6 US-10-411-910B-213	Sequence 213, App
C 22	111.5	2.7	91924	6 US-10-746-294A-48	Sequence 48, Appl
23	110.5	2.7	1805	1 PCT-US04-31524-92	Sequence 92, Appl
C 24	110	2.7	77040	7 US-60-612-215-1	Sequence 1, Appli
25	108	2.7	77040	7 US-60-612-215-1	Sequence 1, Appli
C 26	107	2.6	140064	6 US-10-746-294A-47	Sequence 47, Appl
C 27	106	2.6	10120	7 US-60-613-292-991	Sequence 991, App
28	104.5	2.6	5022	7 US-60-612-215-3	Sequence 3, Appli
C 29	104	2.6	2535	6 US-10-411-910B-244	Sequence 244, App
30	103.5	2.5	8233	6 US-10-474-148-29	Sequence 29, Appl
31	103	2.5	1129	6 US-10-950-009-370	Sequence 370, App
32	101.5	2.5	487	6 US-10-948-737-6292	Sequence 6292, Ap
C 33	101	2.5	1746	6 US-10-411-910B-223	Sequence 223, App
34	101	2.5	2627	7 US-60-613-292-514	Sequence 514, App
35	100	2.5	934	6 US-10-950-009-125	Sequence 125, App
36	100	2.5	4224	1 PCT-US04-14421-52	Sequence 52, Appl
C 37	99.5	2.4	84233	6 US-10-746-294A-105	Sequence 105, App
38	98.5	2.4	533	6 US-10-950-009-371	Sequence 371, App
39	98.5	2.4	1719	7 US-60-612-215-22	Sequence 22, Appl
C 40	98.5	2.4	3091	6 US-10-488-197-19	Sequence 19, Appl
C 41	97.5	2.4	1563	6 US-10-411-910B-195	Sequence 195, App
42	97.5	2.4	4403	1 PCT-US04-14421-76	Sequence 76, Appl
43	97.5	2.4	5034	7 US-60-612-215-4	Sequence 4, Appli
C 44	97	2.4	1911	6 US-10-411-910B-225	Sequence 225, App
C 45	96.5	2.4	2100	6 US-10-509-131-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-10-746-294A-96/c
; Sequence 96, Application US/10746294A
; GENERAL INFORMATION:
; APPLICANT: Bush, David F.
; APPLICANT: Rounsley, Steven D.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Plant Polymorphic Markers and Uses Thereof
; FILE REFERENCE: 38-10(15493)E
; CURRENT APPLICATION NUMBER: US/10/746,294A
; CURRENT FILING DATE: 2003-12-29
; PRIOR APPLICATION NUMBER: US 09/692,412
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: US 09/534,859
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 09/803,736
; PRIOR FILING DATE: 2001-03-12
; NUMBER OF SEQ ID NOS: 124
; SEQ ID NO 96
; LENGTH: 94349
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-746-294A-96

Alignment Scores:
Pred. No.: 7.66e-187 Length: 94349
Score: 2117.00 Matches: 463
Percent Similarity: 66.02% Conservative: 114
Best Local Similarity: 52.97% Mismatches: 147
Query Match: 51.95% Indels: 153
DB: 6 Gaps: 17

US-10-051-909-32 (1-800) x US-10-746-294A-96 (1-94349)

QY 51 IleLeuGluAspLysMetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsn 70

Qy 667 aSerLeuValIleLeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLe 687
Db 51159 GTCGCTGTAAACACTGGTGATAGAGCTTAGTGAATCTTGGAGTTCAATAACCGGTT 51100
Qy 687 uLeuSerThrValSerValIleValTyrPheCysCysPheValMetGlyPheGlyProIl 707
Db 51099 GATATCGACAGCAAGTGTACGGTGTACCTAAGCTGTTTCGTGATGGGTTTGGGCAAT 51040
Qy 707 eProAsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIl 727
Db 51039 TCCAAACATCCTCTGCTCAGAGATATCCCTACTTCTGTGGCGGTCTCTGCATCACCAT 50980
Qy 727 eCysAlaPheThrPheTrpIleGlyAspIleIleValThrTyrSerLeuProValMetIle 747
Db 50979 ATGTGCCCTCACTTCTGGATCTGTGACATAATCGTCACATTACACTCTCCCAGTCATGCT 50920
Qy 747 uAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeuIleSerPh 767
Db 50919 CAAATCCATGGCATCGCAGGAGTCTTTGGCATTATGCAATCGTCTGTGTCGCGTG 50860
Qy 767 eValPheValPheLeuLysValProGluThrLysGlyMetProLeuGluValIleThrGl 787
Db 50859 GGTTTTGTGTACCTCAAGGTACCAAGACAAAGGAATGCCCTTGAAGTTATCTCTGA 50800
Qy 787 uphePheAlaValGlyAlaLysGlnAlaAlaLysAla 800
Db 50799 GTTCTTCTCGTCGGTGCATAAACAGCAAGACGCTGCAGCT 50760

RESULT 2

US-60-613-292-1068
; Sequence 1068, Application US/60613292
; GENERAL INFORMATION:
; APPLICANT: Gene Logic, Inc
; APPLICANT: HIGGS, Brandon
; APPLICANT: ELASHOFF, Michael
; APPLICANT: MENDRICK, Donna L.
; APPLICANT: PORTER, Mark W.
; APPLICANT: CASTLE, Arthur L.
; APPLICANT: JOHNSON, Kory R.
; TITLE OF INVENTION: MOLECULAR TOXICITY MODELS FROM ISOLATED HEPATOCYTES
; FILE REFERENCE: 044921-5135-PR
; CURRENT APPLICATION NUMBER: US/60/613,292
; CURRENT FILING DATE: 2004-09-28
; NUMBER OF SEQ ID NOS: 1300
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1068
; LENGTH: 2571
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION:
US-60-613-292-1068

Alignment Scores:
Pred. No.: 2,76e-24 Length: 2571
Score: 354.00 Matches: 155
Percent Similarity: 33.41% Conservative: 119
Best Local Similarity: 18.90% Mismatches: 205
Query Match: 8.69% Indels: 341
DB: 7 Gaps: 21

US-10-051-909-32 (1-800) x US-60-613-292-1068 (1-2571)

Qy 4 GlySerTrpLeuAlaValGlnThrProPheThrProAspLeuAspArgGluArgLeu 23
Db 133 GGGAGACGCATAGTCACAGAACGTCCATTCTCC----- 165
Qy 24 LeuProSerValValLeuAlaLeuProGlyProLeuProProAlaSer-CysSerSerGl 43
Db 166 -----GTTTCACAGCCCGCACAGCTTGAGCCTCGAGCGCGCGGCCAT 210
Qy 43 nGluProValThrSerAspIleLeuGluAspLysMetSerGlyAlaValLeuValAl 63

Db 211 GGAGCCCAAGCAAG-----AAGGTGACGGCGCCCTTATGTTGGC 252
Qy 63 aile--ValAlaSerIleGlyAsnLeuLeuGlnGlyTrpAspAsnAlaThrIleAlaAl 82
Db 253 CGTGGAGGGCAGTCTCGGATCCCTGCAGTTCGGCTATAACACCGGTGTCAACACGC 312
Qy 82 aAlaValLeuTyrIleLysLysGluPhe--GlnLeuGlnAsnGlu----- 96
Db 313 CCCCCAGAGGTAATTGAGGAGTTCTACAATCAAAACATGGAACACCGCTATGGAGAGTC 372
Qy 97 -----ProThrValGluGlyLeuIleValSerMetSerLeuIleGl 110
Db 373 CATCCCATCCACCACTCAACACACTCTGCTCTCTCCGTGGCCATCTTCTCTGTGCGG 432
Qy 110 yAlaThrIleValThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMe 130
Db 433 GGGCATGATTGTTCTCTCTGTGGGCTCTTTGTTAATCGCTTTGGCAGGCGGAATC 492
Qy 130 tLeuIleLeuSerSerIleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSer---- 148
Db 493 CATGCTGATGATGAACCTGTTGGCCTTTGTGTCTGCCGTGCTTATGGGTTTCTCCAACT 552
Qy 149 ----ProAsnValTyrValLeuLeuAlaArgPheValAspGlyPheGlyIleGlyLe 167
Db 553 GGGCAAGTCCCTTGGATGCTGATCCTGGGCGCTTCATCATTTGGAGTGTACTGTGCCT 612
Qy 167 uAlaValThrLeuValProLeuTyrIleSerGluIleAlaProSerGluIleArgGlyLe 187
Db 613 GACCACCGGCTTTGTGCCCATGTATGTGGGGAGGTGTACCCACACAGCTCTTCGTGGAGC 672
Qy 187 uLeuAsnThrLeuProGlnPheSerGlySerGlyMetPheLeuSerTyrCysMetVa 207
Db 673 CCTGGGCACCCCTGACCAAGTGGCATCGTGGTGGGATCCTTATTGCC-----CAGGT 726
Qy 207 lPheGlyMet---SerLeuSerProSerProAsp---TrpArgIleMetLeuGlyValLe 225
Db 727 GTTCGGCTTAGACTCCATCATGGCAATGCAGACTTGTGGCCTCTACTGTCAGTGTCA 786
Qy 225 uAlaIleProSerLeuPhePheGlyLeuThrIlePheTyrLeuProGluSerProAr 245
Db 787 CTTTCATCCAGCCCTGCTACAGTGT---ATCCTGTTGCCCTTCTGCCCTGAGAGCCCCCG 843
Qy 245 gTrp---LeuValSerLysGlyArgMetAlaGluAlaLysLysValLeuGlnLysLeuAr 264
Db 844 CTTCTGCTCATCAATCGTAACGAGGAGAACCCGGCCCAAGAGTGTGCTGAAAAAGCTTCG 903
Qy 264 gGlyLysAspAspValSerGlyGluLeuSerLeuLeuGluGlyLeuGluValGlyGl 284
Db 904 AGGACAGCCGATGTGACCCGAGACCTGCAGGAGATGAAAGAA----- 946
Qy 284 yAspThrSerIleGluGluTyrIleIleGlyProAlaThrGluAlaAlaAspLeuVa 304
Db 946 ----- 946
Qy 304 lThrAspGlyAspLysGluGlnIleThrLeuTyrGlyProGluGlyGlnSerTrpIl 324
Db 946 ----- 946
Qy 324 eAlaArgProSerLysGlyProIleMetLeuGlySerValLeuSerLeuAlaSerArgHi 344
Db 946 ----- 946
Qy 344 sGlySerMetValAsnGlnSerValProLeuMetAspProIleValThrLeuPheGlySe 364
Db 946 ----- 946
Qy 364 rValHisGluAsnMetProGlnAlaGlyGlySerMetArgSerThrLeuPheProAsnPh 384
Db 946 ----- 946
Qy 384 eGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTrpAspGluGluAs 404

Db 946 ----- 946
QY 404 nLeuHisArgAspGluGluTyrAlaSerAspGlyAlaGlyGlyAspTyrGluAspAs 424
Db 946 ----- 946
QY 424 nLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIleValHi 444
Db 947 ----- 946
QY 444 sHisGlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeuGlyGluGlyGl 464
Db 967 GGAG-----AAGAAGGTCAACCATCTTG----- 988
QY ~464 yAspGlyValSerSerThrAspIleGlyGlyGlyTyrGlnLeuAlaTrpLysTrpSerGl 484
Db 988 ----- 988
QY 484 uLysGluGlyGluAsnGlyArgLysGluGlyGlyPhelysArgValTyrLeuHisGlnGl 504
Db 988 ----- 988
QY 504 uGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGlyGlyAspValPh 524
Db 988 ----- 988
QY 524 eGluGlySerGluPheValHisAlaAlaLeuValSerGlnSerAlaLeuPheSerLy 544
Db 988 ----- 988
QY 544 sGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisProSerGluValAlaAl 564
Db 988 ----- 988
QY 564 aLysGlySerArgTrpLysAspLeuPhe---GluProGlyValArgArgAlaLeuLeuVa 583
Db 989 -----GAGCTGTTCCGCTCACCGCCCTACCGCCAGCCCATCTCAT 1029
QY 583 lGlyValGlyIleGlnIleLeuGlnPheAlaGlyIleAsnGlyValLeuTyrTrTh 603
Db 1030 CGCCGTGGTGTGCTGAGCTGTCCAGCAGCTGTCCGSCATCAATGCTGTGTCTACTACTC 1089
QY 603 rProGlnIleLeuGluGlnAlaGlyVal--AlaValIleLeuSerLysPheGlyLeuSe 622
Db 1090 AACGAGCATCTTCGAGAGCAGGAGGTGTGCAGCAGCTGTGTATGCCACCATCGGCTCGG 1149
QY 622 rSerAlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIl 642
Db 1150 TATCGTCAACACGCGCTTCACTGTGTGTGCTG----- 1183
QY 642 eGlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeu---LeuLeuGlyTh 661
Db 1184 -----TTCTGCTGGAGCGAGCTGGCGCTCGGACCTGCATCTCATTTGGTCT 1230
QY 661 rIleProIleLeu---IleAlaSerLeuValIleLeuValValSerAsnLeuIleAspLe 680
Db 1231 GGCTGGCATGGCGGCTGTGCTGTGCTCATGACCATCGCCCTGGCCCTGCTGGAGCAGCT 1290
QY 680 uGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyrPheCysCysPh 700
Db 1291 GCCCTGGATGTCCTAT-----CTGAGTATCTGGCCCATCTTTGGCTTTGTGGCCTTCTT 1344
QY 700 eValMetGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePheProThrArgVa 720
Db 1345 TGAAGTAGGCCCTGGTCTATTCCATGTTTCATGTGGCCGAGCTGTTCAGCCAGGGGCC 1404
QY 720 lArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIleGlyAspIleIleValTh 740
Db 1405 CCGACCTGCTGTGCTGTGGCTGCTTCTTAACCTGAGACCTCAAACTTCATCTGTTGG 1464
QY 740 rTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAl 760
Db 1465 CATGTGCTTCCAATATGTGGAGCAACTGTGTGGC---CCCTACGCTTTCATCTCTCAC 1521

QY 760 aValValCysLeuIleSerPheValPheValPheLeuLysValProGluThrLysGlyMe 780
Db 1522 GGTGCTGCTGTACTTCTTTCATCTTCACTACTTCAAGTTCTGAGACCAAGGCGG 1581
QY 780 tProLeuGluValIleThrGluPhePheAlaValGlyAlaLysGlnAlaAlaLys 799
Db 1582 GACCTTCGATGAGATCGCTTCCGGCTTCCGGCAGGGGGGTGCCAGCAGGACAAG 1639
RESULT 3
US-60-613-292-899
; Sequence 899, Application US/60613292
; GENERAL INFORMATION:
; APPLICANT: Gene Logic, Inc
; APPLICANT: HIGGS, Brandon
; APPLICANT: ELASHOFF, Michael
; APPLICANT: MENDRICK, Donna L.
; APPLICANT: PORTER, Mark W.
; APPLICANT: CASTLE, Arthur L.
; APPLICANT: JOHNSON, Kory R.
; TITLE OF INVENTION: MOLECULAR TOXICITY MODELS FROM ISOLATED HEPATOCYTES
; FILE REFERENCE: 044921-5135-PR
; CURRENT APPLICATION NUMBER: US/60/613,292
; CURRENT FILING DATE: 2004-09-28
; NUMBER OF SEQ ID NOS: 1300
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 899
; LENGTH: 2573
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION:
US-60-613-292-899

Alignment Scores:
Pred. No.: 3.82e-24 Length: 2573
Score: 352.50 Matches: 157
Percent Similarity: 31.68% Conservative: 105
Best Local Similarity: 18.98% Mismatches: 212
Query Match: 8.65% Indels: 353
DB: 7 Gaps: 18

US-10-051-909-32 (1-800) x US-60-613-292-899 (1-2573)
QY 1 IleArgSerGlySerTrpLeuAlaValGlnThrProPheThrProAspLeuAspArgArg 20
Db 65 ATTGGGAACAGGACCTGGATTGAGACAACAACACTCCGCACACACATGTGAGAAGACAAGA 124
QY 21 GluArgLeuLeuProSerValValLeuAlaLeuProGlyProLeuProProAlaSerCys 40
Db 125 -----TCACCGGAACCTTGGCTTTCACCTG 148
QY 41 SerSerGlnGlu-----ProValThrSerAsp--AspIle----- 51
Db 149 TCTTCACTGCAGTGTGGTTCCTTCCAGTTCGGATATGACATCGGTGTGATCAATGCAC 208
QY 52 -----LeuGluAspLys- 55
Db 209 CTCAAGAGGTAATAATATCCCATTCGACATGTTTGGTGTTCCTCTGGATGACCGAA 268
QY 56 -----MetSerGlyAlaValLeuValAlaIleValAlaSerIleG 69
Db 269 GAGCTACCATTAATATGACATCAATGGCAGACACACCCACTCATAGTCACACCAGCAC 328
QY 69 lyAsnLeuLeuGlnGlyTrpAspAsnAlaThrIleAlaAlaValLeuTyrIleLysL 89
Db 329 ATACGACACCAGCGCTGGGAAGAGAGAGTGAAGGATCTGTCTCATATAGTC----- 381
QY 89 ysGluPheGlnLeuGlnAsnGluProThrValGluGlyLeuIleValSerMetSerLeuI 109
Db 382 -----ACTATGCTGTGGTCTCTGTCTGTGTCCAGCTTTGCAG 418
QY 109 leGlyAlaThrIleValThrThrPheSerGlyProLeuSerAspSerIleGlyArgArgp 129

; APPLICANT: HANSEN, RHONDA
; APPLICANT: KAUFMANN, JOERG
; APPLICANT: KENNEDY, GIULIA C.
; APPLICANT: LAMSON, GEORGE
; APPLICANT: MOLER, EDWARD J.
; APPLICANT: RANDAZZO, FILIPPO
; APPLICANT: REINHARD, CHRISTOPH
; APPLICANT: SUDDUTH-KLINGER, JULIE
; TITLE OF INVENTION: GENE PRODUCTS DIFFERENTIALLY EXPRESSED
; TITLE OF INVENTION: IN CANCEROUS CELLS III
; FILE REFERENCE: 2300-21987
; CURRENT APPLICATION NUMBER: US/10/948,737
; CURRENT FILING DATE: 2004-09-22
; PRIOR APPLICATION NUMBER: 10/616,900
; PRIOR FILING DATE: 2003-07-09
; PRIOR APPLICATION NUMBER: 09/872,850
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: 60/208,871
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 10/081,519
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: 60/270,959
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 10/310,673
; PRIOR FILING DATE: 2003-01-08
; PRIOR APPLICATION NUMBER: 60/336,613
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: US03/00657
; PRIOR FILING DATE: 2003-01-08
; PRIOR APPLICATION NUMBER: 60/345,637
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: 10/081,124
; PRIOR FILING DATE: 2002-02-21
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 13996
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10147
; LENGTH: 582
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-948-737-10147

Alignment Scores:
Pred. No.: 1.06e-09 Length: 582
Score: 187.50 Matches: 38
Percent Similarity: 55.56% Conservative: 22
Best Local Similarity: 35.19% Mismatches: 45
Query Match: 4.60% Indels: 3
DB: 6 Gaps: 1

US-10-051-909-32 (1-800) x US-10-948-737-10147 (1-582)
QY 687 LeuLeuSerThrValSerValIleValTyrPheCysPheValMetGlyPheGlyPro 706
Db 16 CTGCTGGCCACC-----ATGCTCTTCATCATGCGGCTACGCCGTGGGCTGCC 66
QY 707 IleProAsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIleAla 726
Db 67 ATCACCCTGGCTCATGTCTGAGGTCTGCCCCCTGCGTGGCGTGGCGTGGCCTCAGGG 126
QY 727 IleCysAlaPheThrPheTrpIleGlyAspIleIleValThrTyrSerLeuProValMet 746
Db 127 CTCTGCGTGTGGCCAGCTGGCTACCGCCTTCCTCCTCACCAAGTCTCTCTGCCAGTG 186
QY 747 LeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeuIleSer 766
Db 187 GTGAGCACCTTCGGCCCTCCAGGTGCCTTCTCTCTTCCTGCGGCCATCTGCTGGTGAGC 246
QY 767 PheValPheValPheLeuLysValProGluThrLysGlyMetProLeuGluValIleThr 786
Db 247 CTGGTGTTCACAGGCTGTGTGTGCCCGGAGACCAAGGACGGTCCCTGGAGCAGATCGAG 306
QY 787 GluPhePheAlaValGlyAlaLys 794

Db 307 TCCTTCTTCGCGCAGGGGAGAGG 330
RESULT 5
US-60-613-292-497
; Sequence 497, Application US/60613292
; GENERAL INFORMATION:
; APPLICANT: Gene Logic, Inc
; APPLICANT: HIGGS, Brandon
; APPLICANT: ELASHOFF, Michael
; APPLICANT: MENDRICK, Donna L.
; APPLICANT: PORTER, Mark W.
; APPLICANT: CASTLE, Arthur L.
; APPLICANT: JOHNSON, Kory R.
; TITLE OF INVENTION: MOLECULAR TOXICITY MODELS FROM ISOLATED HEPATOCYTES
; FILE REFERENCE: 044921-5135-PR
; CURRENT APPLICATION NUMBER: US/60/613,292
; CURRENT FILING DATE: 2004-09-28
; NUMBER OF SEQ ID NOS: 1300
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 497
; LENGTH: 1910
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION:
US-60-613-292-497

Alignment Scores:
Pred. No.: 2.69e-08 Length: 1910
Score: 181.00 Matches: 75
Percent Similarity: 44.97% Conservative: 59
Best Local Similarity: 25.17% Mismatches: 98
Query Match: 4.44% Indels: 66
DB: 7 Gaps: 15

US-10-051-909-32 (1-800) x US-60-613-292-497 (1-1910)

QY 23 LeuLeuProSerValValLeuAlaLeuProGly-----ProLeuProPro 37
Db 179 CTCTGCGCGTCTTCATGCGCGCTGTGCCTGCCACCATTGTGCCCTGCCCGTGCCT 238
QY 38 AlaSerCysSerSerGlnGlu-----ProValThrSerAsp 50
Db 239 GCCAACCTCAGTCACCAAGGACTTATGGCTGGAAGCCCATCTACCCGGGAGACTGAC--- 295
QY 51 IleLeuGluAspLysMetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsn 70
Db 296 -----GGCAGCTTTAGCTCCTGCTCCGATTTGCCCTATCCCCAGACTGTCCCAAT 346
QY 71 LeuLeu----- 72
Db 347 GTCACTTTGGGACAGAGGTGTCCAACCTCTGGGAGCCTGAGGGTGAGCCCTCACGGTG 406
QY 73 -----GlnGlyTrpAsp-----AsnAlaThrIleAlaAla 83
Db 407 CCTGCTCTCAGGGCTGGGAGTACGACCGCTCAGAATTCTCCTCCACCATTTGCAACTGAG 466
QY 84 ValLeuTyrIleLysLysGluPheGlnLeuGlnAsnGluProThrValGluGlyLeuIle 103
Db 467 TGGCATCTTGTGTGTGTCAGCAGAGAGGACTG---AACAAATATTACGTCCACCTGCTTCTTC 523
QY 104 ValSerMetSerLeuIleGlyAlaThrIleValThrThrPheSerGlyProLeuSerAsp 123
Db 524 ATTGGTGTG---CTGGTGGGAGCCGTGGTGTAT-----GGATACTTGTCTGAC 568
QY 124 SerIleGlyArgArgPrometLeuIleLeuSerSerIleLeuTyrPhePheSerGlyLeu 143
Db 569 AGGTTTGGCAGGCGCGGCTTCTGTGTTGGTGGCTACGTGAGCTCCCTGGTGGTCTG 628
QY 144 IleMetLeuTyrProAsnValTyrValLeuLeuAlaArgPheValAspGlyPhe 163
Db 629 ATGTCTGCAGCCTCCATCACTACATCATGTTCGTAGTCACCCGTACACTCACCGGCTCA 688

; TITLE OF INVENTION: MOLECULAR TOXICITY MODELS FROM ISOLATED HEPATOCYTES
; FILE REFERENCE: 044921-5135-PR
; CURRENT APPLICATION NUMBER: US/60/613,292
; CURRENT FILING DATE: 2004-09-28
; NUMBER OF SEQ ID NOS: 1300
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1274
; LENGTH: 1882
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION:
US-60-613-292-1274

Alignment Scores:
Pred.No.: 3.74e-06 Length: 1882
Score: 158.00 Matches: 123
Percent Similarity: 36.12% Conservative: 89
Best Local Similarity: 20.95% Mismatches: 209
Query Match: 3.88% Indels: 167
DB: 7 Gaps: 28

US-10-051-909-32 (1-800) x US-60-613-292-1274 (1-1882)
Qy 10 GlnThrProPheThrProAspLeuArgArg-----GluArgLeu 23
Db 188 CAGAATCCTGGGGTGGCTGAGCTGAGCCAGCGGTGGTGGAGCCAGGAGGAGCTG 247
Qy 24 -----LeuProSerValValLeuAlaLeuProGlyProLeuProAlaSer--- 39
Db 248 AACTACACTGTCCCGGCGCTG-----GGACCTTCGGACGAGGCTCCTTC 292
Qy 40 -----CysSerSerGlnGluProValThrSerAspAspIleLeuGlu-----Asp 54
Db 293 CTCAGCCAGTGCATGAGGTATGAGGTGAGCTGGAACCCAGAGCACCCCTTGACTGTGTGGAC 352
Qy 55 LysMetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuGlnGly 74
Db 353 CCACTGTCCAGCCTGGTTGCCAACAGGAGTCAAGTGCATTTGGGCCCTCGGAGCATGGC 412
Qy 75 TrpAspAsnAlaThrIleAlaAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGln 94
Db 413 TGGGTATACGACACTCCCGGCTCCTCCATCGTC-----ACTGAGTTTAACCTGGTG 463
Qy 95 ---AsnGluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIle 113
Db 464 TGTGGAGACGCTGGAAGTGGACCTTTTTCAGTCTGTGTGAACITGGGCTTCTTCCTG 523
Qy 114 ValThrThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeu 133
Db 524 GGCTCCCTGGTTGTGGGTACATTGCAGACAGGTTTGGCCGTAAGCTCTGTCTCTGGTG 583
Qy 134 SerSerIleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyrVal 153
Db 584 ACCACGCTGGTCACATCTGTGTCCGGTGTGCTAACAGCGGTGGCCCCAGACTATACATCC 643
Qy 154 LeuLeuLeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValPro 173
Db 644 ATGTTGCTCTTTCCGCTCTGTCAGGGCATG----- 673
Qy 174 LeuTyrIleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGln 193
Db 674 -----GTCAGCAAG-----GGCAGCTGGGTGTCCGGC---TATACCTTGATCACAGAG 718
Qy 194 PheSerGlySerGly-----GlyMetPheLeuSerTyrCysMetValPheGlyMet 210
Db 719 TTTGTGGCTCTGGCTACAGGAGAACGACGGCCATTTGTACCAGATGGCCTTCACAGTG 778
Qy 211 SerLeu-----SerProSerProAspTrpArgIleMetLeuGly 223
Db 779 GGGCTAGTGGGGCTTGGCGGGGTGGCCTATGCCATTCAGACTGGCGCTGGCTCCAGCTA 838
Qy 224 ValLeuAlaIleProSerLeuPhePheGlyLeuThrIlePheTyrLeuProGluSer 243

Db 839 GCTGTGTCCTGCCTACC---TTCCTCTTCTGCTGTTATTACTGGTTTGTCCAGAAATCC 895
Qy 244 ProArgTrpLeuValSerLysGlyArgMetAlaGluAlaLysLysValLeuGlnLysLeu 263
Db 896 CCGCGTGGCTGTTGTCCAGAAAGAGAACACCGCGAGCTGTCAAGATAATGGAGCAAAAT 955
Qy 264 ArgGlyLysAspAspValSerGlyGluLeuSerLeuLeuGluGlyLeuGluValGly 283
Db 956 GCACAGAAGAACGGGAAGTGCCTCCTGCTGACCTGAAGATGCTCTGCCTTGAGGA-GGA 1014
Qy 284 GlyAspThrSerIleGluGluTyrIleIleGlyProAlaThrGluAlaAlaAspLeu 303
Db 1015 TGCCTCAGAAAAGCGAAGTCTGTTTGGCCGACCTGTTCGCACTCCCAA----- 1065
Qy 304 ValThrAspGlyAspLysGluGlnIleThrLeuTyrGlyProGluGluGlyGlnSerTrp 323
Db 1066 -----CCTGAGGAAGCACAC----- 1080
Qy 324 IleAlaArgProSerLysGlyProIleMetLeuGlySerValLeuSerLeuAlaSerArg 343
Db 1081 ---CGTCATCCTGATGATATCTATGTTCTCTTGTGTGTGTGTACCA---GGGTCTCAT 1134
Qy 344 HisGlySerMetValAsnGlnSerValProLeuMetAspProIleValThrLeuPheGly 363
Db 1135 CATGCACGTGGGAGCCACAGGGGCCAACCTCTA-----CCTGGACTTCTTTATTCTTC 1188
Qy 364 Ser-----ValHisGlu-----Asn 368
Db 1189 TCTGGTGAATTTCCCGCGGCTTTCATCATCTCTGTCACCATTTGACCGCATGGCCGCAT 1248
Qy 369 MetProGlnAla-----GlyGlySerMet---ArgSerThrLeuPhe 381
Db 1249 CTACCCAATAGCGGCTCGAATCTGGTGACGGGGGACGCTGCCTCCTCATGATCTTTAT 1308
Qy 382 ProAsnPheGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTrpAsp 401
Db 1309 CCGCATGAGCTGCACTGGTT-----GAACGTTACCCT 1341
Qy 402 GluGluAsnLeuHisArgAspAspGluGluTyrAlaSerAspGlyAlaGlyGlyAsp--- 420
Db 1342 CCGCTGTCTTGGCCGTATGGGGGCCACCATTTGCTGCTGCAGATGGTCTGCTGGTGAACGC 1401
Qy 421 -----TyrGluAsp----- 423
Db 1402 TGAGCTGTACCTACATTCATCAGGAATCTTGGGATGATGGTATGCTGCTGCCCTGTGTGA 1461
Qy 424 -----AsnLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLys 440
Db 1462 CTGGGTGGATCTTCAC---CCCTTTCATGTTGTTTTCAGGCTGATGGAAGTTTGGCAAGC 1518
Qy 441 AspIleValHisHis-----GlyHisArgGlySerAla 451
Db 1519 CTGCCCCCTCATTTTGTGGGTGGTGGGCTGACTGCTGGGCCAT----- 1566
Qy 452 LeuSerMetArgArgGlnThrLeuLeuGlyGlyGlyAspGlyValSerSerThrAsp 471
Db 1567 -----GACTCTTCTTCTCCAGAGAC 1587
Qy 472 IleGlyGlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArg 491
Db 1588 CAAGGGTGTGGCTTTCCTGAGACTATTGAAGAAGCAGAGAACCTGGGGAG----- 1638
Qy 492 LysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArg 511
Db 1639 -----GAGGAATCAAAAGCCCAAGAAAA 1662
Qy 512 GlySerIleValSerLeuPro 518
Db 1663 CACGATTACCTTCAGGTCCA 1683

Db 103 TTGACTAGGACCGAGAGTGTAAGCCAGAACCGGAGATAACC-----GCTCCG 150
Qy 37 ProAlaSerCysSerSerGlnGluProValThrSerAspIleLeuGluAspLysMet 56
Db 151 CCTCACTACGCTTTTCCCGTTCTTTCAAGACA----- 183
Qy 57 SerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrpAsp 76
Db 184 -----GTTTAAATAGCTCAGTGGCTTTCACTGGGTTTTTC----- 219
Qy 77 AsnAlaThrIleAlaAlaValLeuTyr-----IleLysLysGluPhe 91
Db 220 ---TCCACAATAGCAGGTGCCATCTACTATCCAGTTCTGAGCGTTATAGAAAGAAATTC 276
Qy 92 GlnLeuGlnAsnGluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAla 111
Db 277 GATATTGACGAGAA-----TTGGTGAATGTC 303
Qy 112 ThrIleValThr-----ThrPheSerGlyProLeuSer 122
Db 304 ACTGTTGTAGTATATTTTGTATTTCAGGGTCTTGCCGCCACATTCATGGCGGGTTTGC 363
Qy 123 AspSerIleGlyArgArgProMetLeuIleLeuSerSerIleLeuTyrPhePheSerGly 142
Db 364 GATTCAGTGGCAGGAGACCGGTGGTGTCTGTCGCAATCGTCATTATTGTTGCTGC 423
Qy 143 LeuIleMetLeuTrpSerProAsnValTyrValLeuLeuAlaArgPheValAspGly 162
Db 424 ATCGGTCTTGCTTGTGCTCAAAAGTATGCTCAGATCATTTGTGCTAAGATGTCTACAAGCC 483
Qy 163 PheGlyIleGlyLeuAlaValThrLeuValProLeuTyrIleSerGluIleAlaPro--- 181
Db 484 GCCGTATTTCACCCGTGATTCGATTAACAGCGGAATAATGGAGATGTACTACTAGA 543
Qy 182 SerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSerGlySerGlyMetPhe 201
Db 544 GCCGAGCGCGCGGTACGTGGATATGTTGCTGGATTTCAAGTTCTAGGTTCTGCGTTC 603
Qy 202 LeuSerTyrCysMetValPheGlyMetSerLeuSerProSerProAspTrpArgIleMet 221
Db 604 GGAGCC-----CTTATCGGTGCCGATTCATCTAGATGGGGATGGAGAGCCATC 654
Qy 222 LeuGlyValLeuAlaIleProSerLeuPhePheGlyLeuThrIlePheTyrLeuPro 241
Db 655 TTTTGGTCTTAGCAATTGGATCTGGCATTGCTTCCTAGCCTCGTTTTTAATTTTGCCA 714
Qy 242 GluSerProArgTrpLeuValSerLysGly 251
Db 715 GAAACAAAGAGGAACATATCCGGGAATGGT 744

RESULT 10

US-10-948-737-5817/c
; Sequence 5817, Application US/10948737
; GENERAL INFORMATION:
; APPLICANT: CHAN, VIVIEN W.
; APPLICANT: ESCOBEDO, JAIME
; APPLICANT: GARCIA, PABLO DOMINGUEZ
; APPLICANT: HANSEN, RHONDA
; APPLICANT: KAUFMANN, JOERG
; APPLICANT: KENNEDY, GIULIA C.
; APPLICANT: LAMSON, GEORGE
; APPLICANT: MOLER, EDWARD J.
; APPLICANT: RANDAZZO, FILIPPO
; APPLICANT: REINHARD, CHRISTOPH
; APPLICANT: SUDDUTH-KLINGER, JULIE
; TITLE OF INVENTION: GENE PRODUCTS DIFFERENTIALLY EXPRESSED
; TITLE OF INVENTION: IN CANCEROUS CELLS III
; FILE REFERENCE: 2300-21987
; CURRENT APPLICATION NUMBER: US/10/948,737
; CURRENT FILING DATE: 2004-09-22
; PRIOR APPLICATION NUMBER: 10/616,900
; PRIOR FILING DATE: 2003-07-09
; PRIOR APPLICATION NUMBER: 09/872,850

; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: 60/208,871
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 10/081,519
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: 60/270,959
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 10/310,673
; PRIOR FILING DATE: 2003-01-08
; PRIOR APPLICATION NUMBER: 60/336,613
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: US03/00657
; PRIOR FILING DATE: 2003-01-08
; PRIOR APPLICATION NUMBER: 60/345,637
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: 10/081,124
; PRIOR FILING DATE: 2002-02-21
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 13996
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5817
; LENGTH: 411
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-948-737-5817

Alignment Scores:
Pred. No.: 1.24e-05 Length: 411
Score: 141.50 Matches: 35
Percent Similarity: 56.78% Conservative: 32
Best Local Similarity: 29.66% Mismatches: 44
Query Match: 3.47% Indels: 7
DB: 6 Gaps: 4

US-10-051-909-32 (1-800) x US-10-948-737-5817 (1-411)
Qy 118 SerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSerIleLeu 137
Db 348 ACTGGCCTGACCTGTGACCCCGTCCCCAGCGGGAATCAATGCTGATGATGAACCTGCCG 289
Qy 138 TyrPhePheSerGlyLeuIleMetLeuTrpSer-----ProAsnValTyrValLeu 154
Db 288 GCCTTCGTGTCGCCGCTCATCGGCTTCTCGAAACTGGCAAGTCCTTTGAGATGCTG 229
Qy 155 LeuLeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeu 174
Db 228 ATCCTGGGCGCTTTCATCATCGGTGTGTACTGCGGCTGACCACAGGCTTCGTGCCCATG 169
Qy 175 TyrIleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPhe 194
Db 168 TATGTGGTGAAGTGTACCCACACAGCCCTTCGTGGGGCCCTGGGCACCTGCACAGCTG 109
Qy 195 SerGlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMet---SerLeuSer 213
Db 108 GGCATCGTCTCGGCATCCTCATCGCC-----CAGGTGTTTCGACCTGGACTCCATCATG 55
Qy 214 ProSerProAsp---TrpArgIleMetLeuGlyValLeuAlaIleProSerLeu 230
Db 54 GCAACAAGGACCTGTGGCCCTGCTGCTGAGCATCATCTTTCATCCCGCGCGTA 1

RESULT 11

US-60-613-292-399
; Sequence 399, Application US/60613292
; GENERAL INFORMATION:
; APPLICANT: Gene Logic, Inc
; APPLICANT: HIGGS, Brandon
; APPLICANT: ELASHOFF, Michael
; APPLICANT: MENDRICK, Donna L.
; APPLICANT: PORTER, Mark W.
; APPLICANT: CASTLE, Arthur L.
; APPLICANT: JOHNSON, Kory R.
; TITLE OF INVENTION: MOLECULAR TOXICITY MODELS FROM ISOLATED HEPATOCYTES
; FILE REFERENCE: 044921-5135-PR

```

; CURRENT APPLICATION NUMBER: US/60/613,292
; CURRENT FILING DATE: 2004-09-28
; NUMBER OF SEQ ID NOS: 1300
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 399
; LENGTH: 9888
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION:
US-60-613-292-399

Alignment Scores:
Pred. No.: 0.048 Length: 9888
Score: 126.00 Matches: 154
Percent Similarity: 34.25% Conservative: 120
Best local Similarity: 19.25% Mismatches: 301
Query Match: 3.09% Indels: 226
DB: 7 Gaps: 30

US-10-051-909-32 (1-800) x US-60-613-292-399 (1-9888)
QY 1 IleArgSerGlySerTrpLeuAlaValGlnThrProPheThrProAspLeuAspArgArg 20
Db 498 TTGCGTGCAGCTCTGTGGAGGTTTGCTGAG-----CTGGCTCACCTGGTTCGACCT 548
QY 21 GluArgLeuLeuProSerValValLeuAlaLeuProGlyProLeuProProAlaSerCys 40
Db 549 CAGAAAGTGCAGGCCTTATCTGGTGAATCTTCTTCCA-----TGT 587
QY 41 SerSerGlnGluProValThrSerAspAspIleLeuGluAspLysMetSerGlyAlaVal 60
Db 588 TTGACCCGGAACAAGCAACGACCGGAGGAGTCAGTTCAGGAGACTTTGGCTGCAGCTGT 647
QY 61 LeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrpAspAsnAlaThrIle 80
Db 648 ---CCTAAATATGGCCTCTTTTGGCAAT----- 674
QY 81 AlaAlaAlaValLeuTyriLeLysLysGluPheGlnLeuGlnAsnGluProThrVal--- 99
Db 675 -----TTGCGGAATGACAATGAATTAAGTTCTA 704
QY 100 ---GluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThrThrPheSer 118
Db 705 TTGAAAGCTTTCATAGCAAAATCTGAAGTCAAGCTCTCCACCTGTGCGGCGGACAGCCT 764
QY 119 GlyProLeuSerAspSerIle-----GlyArgArgProMetLeuIleLeuSerSer 135
Db 765 GGGTCA--GCAGTGAGTATCTGCCAGCACTCTAGGAGGACACAGTACTTCTACAACCTGG 821
QY 136 IleLeuTyriPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyriValLeuLeu 155
Db 822 CTCCTGAATGTGCTCCTAGGTTTGTGTTGCCATGGAGGAGCAACACCCCACTCTCCTG 881
QY 156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThr-----LeuVal 172
Db 882 ATC-----CTTGGTGTGTGCTCAGACTGAGGTGTCTAGTG 917
QY 173 ProLeuTyriLeSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuPro 192
Db 918 CCCTTGCTCCAGCAGCAGGTCAAGGACACAAGTCTAAAGGGCAGCTTTGGGGTAAC-ACG 976
QY 193 GlnPheSerGlySerGlyMetPheLeuSerTyriCys-MetValPheGlyMetSerLe 212
Db 977 GAAAGAAATGGAAGTCTCTCCTTCTGCAGACAGCTTGTCCAGGTTTATGAAGTACTT 1036
QY 212 u-----SerProSerProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLe 230
Db 1037 GCATCACACACACCAAGACCAATAATGGTGACAGGGGCAATTGGAGTCTCCTGCAGCA 1096
QY 230 uphePhePheGlyLeuThrIlePheTyriLeuProGluSerProArgTrpLeuValSerLy 250
Db 1097 GCTCTTCCTGACCCCTCCA-----CCTGAGCTGTGTCAGCACTGACCAACC 1144

250 sGlyArgMetAlaGluAlaLysLysValLeuGlnLysLeuArgGlyLysAspValSe 270
Db 1145 AGGAGGGCTCGGCAGCTCACTCTGTGTCGAGAGGAAGCGGGGGCCGAGGCCGC---AG 1201
QY 270 rGlyGluLeuSerLeuLeuGluGly----- 279
Db 1202 CGGAGTATCGTGAGCTTTTAGCTGGAGGGGTTCTCTCATGCGCCCTGTTCTCTCAAG 1261
QY 279 ----- 279
Db 1262 AAAGCAAAAGGAAAGTGCTCTTAGGAGAGGAAGCCCTTGGAGGATGACTCGGAGTC 1321
QY 280 -----LeuGluValGlyGlyAs 285
Db 1322 CAGGTCAGATGTCAGCAGCTCAGCCCTTTCAGCCTCTGTGAAGAGTGAGATTGGTGAGA 1381
QY 285 pThrSerIleGluGluTyriIleIleGlyProAlaThrGluAlaAlaAspLeuValTh 305
Db 1382 GCTCGTGTCTTCTTCTCGGTGTCTCCACTCCCGGTTCTGTAGGTACGACATCATCAC 1441
QY 305 rAspGlyAspLysGluGlnIleThrLeuTyriGlyProGluGluGlyGlnSer-TripIleA 325
Db 1442 TGAGCAGCCTCGATCCAGCACACACTT-----CAAGCAGACTCTGTGGATT 1489
QY 325 laArgProSerLysGlyProIleMetLeuGlySerValLeuSerLeuAlaSer--ArgHi 344
Db 1490 GTCAGGCTGTGACTTGACCACTGCTGCTACTGATGGAGATGAGGAAGACATCTTGAGCCA 1549
QY 344 sGlySerMetValAsnGlnSerValProLeuMetAspProIleValThrLeuPheGlySe 364
Db 1550 CAGCTCCAGCCAGTTCAGTGCTGTTCCA--TCCGACCTGCCATGGACCTGAATGATGG 1606
QY 364 rValHisGluAsnMetProGlnAlaGlyGlySerMetArgSerThrLeuPheProAsnPh 384
Db 1607 GACCCAGGCCTCTCACCCATCAGTGACAGTCTCAGACCACCACTGAGGACCT---- 1661
QY 384 eGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTrpAspGluGluAs 404
Db 1662 -----GATTCAGCTGT 1672
QY 404 nLeuHisArgAspAspGluGluTyriAlaSerAspGlyAlaGlyGlyAspTyri----- 421
Db 1673 GACTCCTTCTGACAGTCTCTGAAATTTGCTTAGATGGTGTGCTGACAGCCAGTATTTAGGCGT 1732
QY 422 -----GluAspAsnLeuHisSerProLeuLeuSerArgGlnAl 434
Db 1733 GCAGATAGGACAGCCACAGGAGGAAGAC-----AGAGAAGC 1768
QY 434 aThrGly-----AlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAl 451
Db 1769 TGCAGGTGTTCTTTCTGTTGTAAGTCTCAGACGTT-----TTCAGAAACTCTTC 1816
QY 451 aLeuSerMetArgArgGlnThrLeuLeuGlyGlyGlyAspGlyValSerSerThrAs 471
Db 1817 TCTGGCCCTTCAGCAGGCACACTTGTGTAAGAAAGTGGTTCATAGCCGCGCAGCCTTCTGA 1876
QY 471 pIleGlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyAr 491
Db 1877 C-----AGCAGTGTGTGATAAGTTGTTTCAAAAGAT----- 1907
QY 491 gLysGluGlyGlyPheLysArgValTyriLeuHisGlnGluGlyValProGlySerArgAr 511
Db 1908 -----GAGGTGCTGAAGTGGGACCCCAAGAAAGCAAGCC 1942
QY 511 gGlySerIleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHi 531
Db 1943 TTGCCGAATCAAAGGTGACATAGGACAGCCCTAATGATGATGATTTCTGCTCTCTGTGTA 2002
QY 531 sAlaAlaAlaLeuValSerGlnSerAlaLeuPheSer----- 543
Db 2003 TTGTGTCCGTCTTTTATCCGCTTCTCTTTTGTAACTGGCGAAAGAAAGCACTGGTTCC 2062
```

QY 544 -----LysGlyLeuAlaGluProArgMetSerAspAlaA1 555
Db 2063 AGACAGAGATGTGAGAGTCAGTGTGAAGCCCTGGCCCTCAGCTGTATTGGTCAGCTGT 2122
QY 555 aMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuPheGluPr 575
Db 2123 GGCCCTTCATCCAGAGTCGTTCTTCAGCAA----- 2153
QY 575 oGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeu-----G1 591
Db 2154 -----CTCTACAAAGTACCTCTCAGTACCATGGAAGTACTGAGGA 2194
QY 591 nGlnPheAlaGly-----IleAsnGlyValLeuTyrThrProGlnIleLeuGlu1 609
Db 2195 ACAGTATGTCTCTGACATCTCGAACTACATCGATCATGGAGACCCCTCAGGTG----- 2246
QY 609 nAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLe 629
Db 2247 -CGAGGAGCTACTGCCATTCTC-----TGTGGGACCCT 2278
QY 629 uIleSerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMe 649
Db 2279 TGCTACTCCATCCTCAGCAGGTCCCGTCTCCGTTGGTGACTGGTGGGCACCATCAG 2338
QY 649 tAspLeuSerGlyArgArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLe 669
Db 2339 GGCCCTGACAGGAAATACATTCTCTGTGGTGGACTGCTTTACTGCAGAAAACCTTT 2398
QY 669 u-----ValIleLeuValValSerAsnLeuIleAspLeu---GlyThrLeuAlaH 685
Db 2399 GAAGGATGAATCTTCTGTACTTGCAGTGGCTTGTACAGCTGTGAGGCACCTGTGTCT 2458
QY 685 isAlaLeuLeuSerThrValSerValIleValTyrPheCysCysPheVal 701
Db 2459 GAGTCTTTGCAGCAGCAGCTACAGTGACTTGGGATTACAACTGCTTATTG 2508

RESULT 12
US-10-950-009-135
; Sequence 135, Application US/10950009
; GENERAL INFORMATION:
; APPLICANT: BERKA, Randy
; APPLICANT: BASHKIROVA, Elena
; APPLICANT: REY, Michael
; TITLE OF INVENTION: Methods For Monitoring Gene Expression
; FILE REFERENCE: 10541.200-US
; CURRENT APPLICATION NUMBER: US/10/950,009
; CURRENT FILING DATE: 2004-09-24
; PRIOR APPLICATION NUMBER: 60/506,140
; PRIOR FILING DATE: 2003-09-25
; NUMBER OF SEQ ID NOS: 1190
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 135
; LENGTH: 473
; TYPE: DNA
; ORGANISM: Trichoderma reesei
US-10-950-009-135

Alignment Scores:
Pred. No.: 0.000748 Length: 473
Score: 123.50 Matches: 45
Percent Similarity: 43.29% Conservative: 26
Best Local Similarity: 27.44% Mismatches: 58
Query Match: 3.03% Indels: 35
DB: 6 Gaps: 6

US-10-051-909-32 (1-800) x US-10-950-009-135 (1-473)

QY 576 GlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnPheAlaGly 595
Db 5 GGTGCGGCGGAGGTACACTTGTGGCGTCTGTCAGATTTGGCAGCAACTCCTCGGT 64
QY 596 IleAsnGlyValLeuTyrThrProGlnIleLeuGluAlaGlyValAlaValIle 615

Db 65 GGTAATGTGATGCTTTACTACCTTGTCTACATCTTCAACATGGCTGGCATG----- 115
QY 616 LeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThrThr 635
Db 116 -----TCCGGAATATACCGCGCTCAGCTCTTCAATTATCAATAT 154
QY 636 LeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArg 655
Db 155 GTCATCTTTCTCGTCACGACTGGAGGAGTTCTCTTTGTGTCGTTGATCGCATTTGGTGGCCGG 214
QY 656 PheLeuLeuLeu-----GlyThrIleProIleLeuIleAlaSerLeu 669
Db 215 TGGTGTCTTATTGTTGGGGCCATCATTTTGGGGCGTCATCCATTTCATCGTCCGTGCC--- 271
QY 670 ValIleLeuValValSerAsnLeuIleAspLeu----- 680
Db 272 GTCATGGCAGTCTACGGGCACCACGTCGACAGTGGATGGAAACGATATTTCTGAGGTGG 331
QY 681 -----GlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyrPhe 697
Db 332 CAGATTGGTGGCCCTCTCTGCCAAGGCCATCATTTGCG-----CTTTGTTACATC 379
QY 698 CysCysPheValMetGly-----PheGlyProIleProAsnIleLeuCysAlaGluIle 715
Db 380 TTTGTTGGAGTCTATGGAGTCACATGGCCCCCGTGCATGGATTACTGCGGAGAAGTC 439
QY 716 PheProThrArg 719
Db 440 TTCCCTTGAAG 451

RESULT 13
US-10-950-009-1142
; Sequence 1142, Application US/10950009
; GENERAL INFORMATION:
; APPLICANT: BERKA, Randy
; APPLICANT: BASHKIROVA, Elena
; APPLICANT: REY, Michael
; TITLE OF INVENTION: Methods For Monitoring Gene Expression
; FILE REFERENCE: 10541.200-US
; CURRENT APPLICATION NUMBER: US/10/950,009
; CURRENT FILING DATE: 2004-09-24
; PRIOR APPLICATION NUMBER: 60/506,140
; PRIOR FILING DATE: 2003-09-25
; NUMBER OF SEQ ID NOS: 1190
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1142
; LENGTH: 491
; TYPE: DNA
; ORGANISM: Trichoderma reesei
US-10-950-009-1142

Alignment Scores:
Pred. No.: 0.00209 Length: 491
Score: 119.00 Matches: 31
Percent Similarity: 50.50% Conservative: 20
Best Local Similarity: 30.69% Mismatches: 46
Query Match: 2.92% Indels: 4
DB: 6 Gaps: 2

US-10-051-909-32 (1-800) x US-10-950-009-1142 (1-491)

QY 700 PheValMetGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePheProThrArg 719
Db 24 TTTGGCTACAGTTGGGACCTTGGCCCTGGATTCTCATCGTGGATCTGGCCCTTGAGC 83
QY 720 ValArgGlyLeuCysIleAlaIleCysAlaPheThrPheThrIleGlyAspIleVal 739
Db 84 TCTCGACCTTATGGTGTCTCTCTGGGAGCTTCCAGCAACTGGATGAACAACATCATCATC 143
QY 740 ThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyr 759
Db 144 GGACAAGTGACTCCGACATGTTGCAAGGATC---ACCTATGGCACCTATATCCTGTTC 200

QY 760 AlaValValCysLeuIleSerPheValPheValPheLeuLysValProGluThrLysGly 779
Db 201 GGAATTATAACATACCTTGGTGGCGCCTTTGTCTACTTCTCGTCCCGAGACCAAGCGT 260
QY 780 MetProLeuGluValIleThrGluPhePheAlaValGlyAlaLysGlnAlaAlaLys 799
Db 261 CTCACCTGGAA-----GAGATGGACATCATCTTTGGCAGCGAGGGCGCGCCCGT 311
QY 800 Ala 800
Db 312 GCC 314
RESULT 14
US-10-632-694A-4/c
; Sequence 4, Application US/10632694A
; GENERAL INFORMATION:
; APPLICANT: Allison, Anthony
; TITLE OF INVENTION: MODIFIED ANNEXIN PROTEINS AND METHODS FOR TREATING VASO-OCCLUSIVE
; TITLE OF INVENTION: SICKLE-CELL DISEASE
; FILE REFERENCE: SURR.113
; CURRENT APPLICATION NUMBER: US/10/632,694A
; CURRENT FILING DATE: 2003-08-01
; PRIOR APPLICATION NUMBER: 60/400,718
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: 10/080,370
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 4
; LENGTH: 2016
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: primer
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (45)..(45)
; OTHER INFORMATION: n = a, c, t, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1000)..(1002)
; OTHER INFORMATION: n = a, c, t, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1051)..(1053)
; OTHER INFORMATION: n = a, c, t, or g
US-10-632-694A-4
Alignment Scores:
Pred. No.: 0.023 Length: 2016
Score: 118.00 Matches: 158
Percent Similarity: 33.87% Conservative: 75
Best Local Similarity: 22.97% Mismatches: 239
Query Match: 2.90% Indels: 218
DB: 6 Gaps: 31
US-10-051-909-32 (1-800) x US-10-632-694A-4 (1-2016)
QY 36 ProProAlaSerCysSerSerGln-----GluProValThrSerAspAspIleLeu 52
Db 1997 CCGCCGCACAGCAGCAGCAGGCGCTTCTTTAGTCGCGGAGGTGTCGCCCTTGATCATG 1938
QY 53 GluAspLysMetSerGlyAlaValLeuValAlaIleValAlaSerIleGly-----Asn 70
Db 1937 GAGTACAGGGA-GGTGGCGAAGTTCTTGGGAACTCCTTGGGAACTCCTTGGGATGTTGAACAGGTGCGAT 1879
QY 71 LeuLeuGlnGlyTrpAspAsnAla-----ThrIleAlaAlaValLeuTyr 86
Db 1878 CTCGGAGCGGGACACGATCAGCGGATCAGGGRGTGGTTCGTGGTGGCGGCGCCCTTCAT 1819
QY 86 ----- 86
Db 1818 GCGGTAGTACAGGGTCTCGGCCAGGTAGGCGGGGATGGAGCGGATGGACTTCACCCAGCG 1759

QY 87 IleLysLysGluPheGlnLeuGlnAsnGluProThrValGluGlyLeuIleValSerMet 106
Db 1758 CAGCAGCAGGTTCTCCAGGTTGCCGAGGTTCTCGGTCGATGTTCTCTCGATCTGGAA 1699
QY 107 SerLeuIleGlyAlaThrIleVal-ThrThrPheSerGlyProLeuSerAspSerIleG1 126
Db 1698 GCCGAGATGTCATGTAATCTTGTGAAACACGGCGCAGGTGGGACACGAGCGGTGCC 1639
QY 126 yArgArgProMetLeuIleLeuSerSerIleLeuTyrPhePheSerGlyLeuIleMetLe 146
Db 1638 CAGGATGGTGAACACTTCTCCTCGTCGGTGCCTTCCCTTCCAGCTCG-----CCGGC 1588
QY 146 uTrpSerProAsnValTyrValLeuLeuAlaArgPheValAspGlyPheGlyIleG1 166
Db 1587 CTGGAACAG-GGCCTGGCGCTCCAGCTCCAGCTCCAGCTGGGCGTCGTCGATGGC--GGTGTCCG 1532
QY 166 yLeuAlaValThrLeu----- 171
Db 1531 GGTCCGGGTTGGCCTGCAGCAGCACCACCAGCATGGCTGGTAGTAGCCGAGGTGTCCG 1472
QY 172 -----ValProLeuTyrIleSerGluIleAlaProSerGluIleArgGlyLeuLeuAs 189
Db 1471 CCACCACGTCGCTCCAGTTGGAGCCGTACTC-----CTCCTCGTAGGCTGCTTGA 1418
QY 189 nThrLeuProGlnPheSerGlySerGlyMetPheLeuSerTyrCysMetValPheG1 209
Db 1417 TGGCGCGCAGCTCCTCGGGGTGCGGAGCGGATGATCTCGGTACACACCTTCTCGTCGG 1358
QY 209 yMetSerLeuSerProSerProAspTrpArgIleMetLeuGlyValLeuAlaIleProSe 229
Db 1357 TGCCGGC-----GCCAGCTTGGCGTGCTTCAGCTCGTAGG 1322
QY 229 rLeuPhePheGlyLeuThrIlePheTyrLeuProGluSerProArgTrpLeuValSe 249
Db 1321 C-----GTCGTACA 1313
QY 249 rLysGlyArgMetAlaGluAlaLysLysValLeuGlnLysLeuArgGlyLys----- 266
Db 1312 GCGGGAGGGCTTTCATCAGGGCCACGATCAGCTTCTGAACCTTGCCTCGGTCAGCTCGGACT 1253
QY 267 -----AspAspValSerGlyGluLeuSerLeuLeuGly----- 279
Db 1252 TCATGTCGTTACACAGTTCGCGGCGGCGGACAGGCTTGAACCTCCTCGGCGATCTGCTGGC 1193
QY 280 -LeuGluValGly---GlyAspThrSerIleGluGluTyrIle----- 292
Db 1192 GCTGGCGTTGGAGCGGCGGTCAGCAGGTTGAGGATGGAGTCTCTCGGTCGCCAGGC 1133
QY 293 -----IleGlyProAlaThrGluAlaAlaAspAspLe 303
Db 1132 CTTTCATGGCCTTGGCAGCAGCCTCGGCGTCGCGGCGGCGCTCGAAGCCGAGAA----- 1078
QY 303 uValThrAspGly-AspLysGluGlnIleThrLeuTyrGlyProGluGluGlyGlnSerT 323
Db 1077 -GTCGGTCACGGTGCCTCGCGCAGGCGNNNAGTACTTCCGGGGCCC-----T 1034
QY 323 rp-IleAlaArgProSerLysGlyProIleMetLeuGlySer-ValLeuSerLeuAlaSe 342
Db 1033 GGAACAGCACCTCCA-----GGCCCGATCGAGATCTNNNGTCCTCCTCGCCCGCACA 980
QY 342 rArgHisGlySerMetValAsnGlnSerValProLeuMetAspProIleValThrLeuPh 362
Db 979 GCAGCAGCAGGCGCTT-----CTTGTAGTCGCCGG 950
QY 362 eGlySerValHisGluAsnMetProGlnAlaGlySerMetArgSerThrLeuPhePr 382
Db 949 AGGTGTGCCCTTGTATCATGGAGTACAGGGAGGTGGCGAAGTCTTTCGGAACCTCTTGC 890
QY 382 oAsnPheGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTrpAspG1 402
Db 889 G----- 889

QY	402	uGluAsnLeuHisArgAspGluGluTyrAlaSerAspGlyAlaGlyGlyAspTyrG1	422
Db	888	-----GATGTTGAACAGGTCGATCTCGAGAGCGGACACGATCAGCG	848
QY	422	uAspAsn-----LeuHisSerProLeuLeuSerArgGlnAl	434
Db	847	GGATCAGGGTGTGTCGTCGGTGC CGCGCCCTTCATGGCGTAGTACAGGGTCTCGGCCA	788
QY	434	aThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeuSerMe	454
Db	787	GGTAGCGCGGATGGAGCGGATGGACTTCACCAACGGC-----	751
QY	454	tArgArgGlnThrLeuLeuGlyGluGlyGly-----AspGlyValSerSerThrAs	471
Db	750	-CAGCAGCAGGTTCTCCAGGTTGCCGAGGTC TC CGCGTCGATGGTCTCCTCGATCTGGA	692
QY	471	pIleGlyGlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluasnglyAr	491
Db	691	AGCCGGAGATGGTCATGTACTTGTCT-----GAACACGCGCGCAGGTGGGACA	644
QY	491	gLYSGluGlyGlyPheLysArgValTyrLeuHisGlnGluGlyValPro-----	507
Db	643	CGGAGCGGGTGCCAGGATGGTGATGAACCTTCTCCTCGTCGGTGCCCCACCTTCAGCTCGC	584
QY	508	-GlySerArgArgGlySerIleValSerLeu-----ProGlyGlyGlyAs	522
Db	583	CGGCCCTGGAACAGGGCCTGGGCGTCCAGCTCCACCTGGCGTCGTCGATGGCGGTGTCGG	524
QY	522	pValPheGluGlySerGluPheValHisAlaAlaLeuValSerGln-----	538
Db	523	GGTCGCGGTTGGCCTGCAGCAGCACACCAGCATCGCTGGTAGTAGCCGAGGTTGTCGC	464
QY	539	-----SerAlaLeuPheSerLysGlyLeuAlaGluProAr	550
Db	463	CCACCAGTCGTCCTCCAGGTGGAGCCGTACTCTCTCGTAGGCCTGCTT-----	412
QY	550	gMetSerAspAlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLy	570
Db	411	-----GATGGCGCGCAGCTC-----CTCGGGGGTGCGGGAGGCGCA	377
QY	570	sAspLeuPheGluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIle--	589
Db	376	TGATCTC-----GGTCAGACCTTCTCGTCGGTGCCGCGCGCCAGCTTGG	332
QY	590	-LeuGlnGlnPheAlaGlyIle-----AsnGlyValLeuTyrTyrThrProGlnI1	606
Db	331	CGTGCTTCAGCTCGTAGGCGTCGTACAGGCGGAGGGCTTCATCAGGGCCACGATCAGCT	272
QY	606	eLeuGluGlnAlaGly	611
Db	271	TCTCGAACTTGCCGGT	256

RESULT 15

US-10-632-694A-5/c
 ; Sequence 5, Application US/10632694A
 ; GENERAL INFORMATION:
 ; APPLICANT: Allison, Anthony
 ; TITLE OF INVENTION: MODIFIED ANNEXIN PROTEINS AND METHODS FOR TREATING VASO-OCCLUSIVE
 ; TITLE OF INVENTION: SICKLE-CELL DISEASE
 ; FILE REFERENCE: SURR.113
 ; CURRENT APPLICATION NUMBER: US/10/632,694A
 ; CURRENT FILING DATE: 2003-08-01
 ; PRIOR APPLICATION NUMBER: 60/400,718
 ; PRIOR FILING DATE: 2002-08-02
 ; PRIOR APPLICATION NUMBER: 10/080,370
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 9
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 5
 ; LENGTH: 2016
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:

Db 1357 TGCCGC-----GCCCCAGCTGGCGTGCTTCAGCTCGTAGG 1322
Qy 229 rLeuPhePheGlyLeuThrIlePheTyrLeuProGluSerProArgTrpLeuValSe 249
Db 1321 C-----GTCTACA 1313
Qy 249 rLysGlyArgMetAlaGluAlaLysLysValLeuGlnLysLeuArgGlyLys----- 266
Db 1312 GCGGAGGGCTTCATCAGGGCCACGATCAGCTTCTCGAACTTCCGGTCAGCTCGGACT 1253
Qy 267 -----AspAspValSerGlyGluLeuSerLeuLeuLeuGluGly----- 279
Db 1252 TCATGTCGTTACCAGGTGCGGCGCCGAACAGGGTCTTGAACTCCTCGCGATCTGTGGC 1193
Qy 280 -LeuGluValGly---GlyAspThrSerIleGluGluTyrIle----- 292
Db 1192 GCTGGCGTTGGAGCGGCGGTACAGAGGTTTACGATGGAGTCTCTCGTGGTCCCCAGGC 1133
Qy 293 -----IleGlyProAlaThrGluAlaAlaAspAspLe 303
Db 1132 CCTTCATGCGCTTGCGCAGCACCTCGCGTGGCGGCGCGTGAAGCCGGAGAA----- 1078
Qy 303 uValThrAspGly-AspLysGluGlnIleThrLeuTyrGlyProGluGluGlyGlnSerT 323
Db 1077 -GTGCGTCACGGTGCCGCGCAGGCGNNNAGTACTTCCGGGGCCC-----T 1034
Qy 323 rp-IleAlaArgProSerLysGlyProIleMetLeuGlySer-ValLeuSerLeuAlaSe 342
Db 1033 GGAACAGCACCTCCA-----GGCCCGATCGAGATCTNNNGTCGTCTCGCCGCGCGACA 980
Qy 342 rArgHisGlySerMetValAsnGlnSerValProLeuMetAspProIleValThrLeuPh 362
Db 979 GCAGCAGCAGGCGCTT-----CTGTAGTCGCCGG 950
Qy 362 eGlySerValHisGluAsnMetProGlnAlaGlyGlySerMetArgSerThrLeuPhePr 382
Db 949 AGGTGTGCGCTTGATCATGGAGTACAGGAGGTGGCGAAGTCTTTCGGAACTCCTTGC 890
Qy 382 oAsnPheGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTrpAspGl 402
Db 889 G----- 889
Qy 402 uGluAsnLeuHisArgAspGluGluTyrAlaSerAspGlyAlaGlyGlyAspTyrGl 422
Db 888 -----GATGTTGAACAGGTCGATCTCGGAGCGGGACACGATCACGC 848
Qy 422 uAspAsn-----LeuHisSerProLeuLeuSerArgGlnAl 434
Db 847 GGATCAGGGTGTGTCGTGCGTGCCGGCGCCCTTCATGGCGTAGTACAGGGTCTCGGCCA 788
Qy 434 aThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeuSerMe 454
Db 787 GGTAGCGGGGATGGAGCGGATGGACTTCACACGGC----- 751
Qy 454 tArgArgGlnThrLeuLeuGlyGluGly-----AspGlyValSerSerThrAs 471
Db 750 -CAGCAGCAGGTTCTCCAGGTTGCCGAGGTCTCGCGTTCGATGGTCTCCTCGATCTGGA 692
Qy 471 pIleGlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyAr 491
Db 691 AGCCGGAGATGGTCATGTACTTGTC-----GAACACGGCGGCAGGTGGGACA 644
Qy 491 gLysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGlyValPro----- 507
Db 643 CGGAGCGGGTGCCAGGATGGTGATGAACCTCTCCTCGTGGTGGCCACTTCAGTCTCGC 584
Qy 508 -GlySerArgArgGlySerIleValSerLeu-----ProGlyGlyGlyAs 522
Db 583 CGGCCTGGAACAGGGCTGGGCGTCCAGCTCCACCTGGGCGTCTCGATGGCGGTGTCGG 524
Qy 522 pValPheGluGlySerGluPheValHisAlaAlaLeuValSerGln----- 538

Db 523 GGTGCGGGTTGGCCTGCAGCAGCACACACAGCATGCGCTGGTAGTAGCCGGAGGTGTCGC 464
Qy 539 -----SerAlaLeuPheSerLysGlyLeuAlaGluProAr 550
Db 463 CCACCACGTCGTCCTCCAGGTTGGAGCCGTACTCCTCCTCGTAGGCTGCTT----- 412
Qy 550 gMetSerAspAlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLy 570
Db 411 -----GATGGCGCGCAGCTC-----CTCGGGGTGCGGGAGGCGA 377
Qy 570 sAspLeuPheGluProGlyValArgArgAlaLeuValGlyValGlyIleGlnIle-- 589
Db 376 TGATCTC-----GGTCAGCACCTTCTCGTCGGTGGCGGCGCCAGCTTGG 332
Qy 590 -LeuGlnGlnPheAlaGlyIle-----AsnGlyValLeuTyrTyrThrProGlnIl 606
Db 331 CGTGCTTCAGCTCGTAGGCGTCTGTACAGCGGGAGGGCTTCATCAGGGCCACGATCAGCT 272
Qy 606 eLeuGluGlnAlaGly 611
Db 271 TCTCGAACTTGCCGGT 256

Search completed: October 13, 2004, 23:34:49
Job time : 333.921 secs

this Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2004, 11:33:22 ; Search time 21.286 Seconds
(without alignments)
3616.147 Million cell updates/sec

Title: US-10-051-909-32
Perfect score: 4075
Sequence: 1 IRSGSWLAVQTPTPDLRR.....PLEVITEFFAVGAKQAAGA 800

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2584	63.4	729	2 T06127	probable sugar tra
2	2278	55.9	734	2 H86340	Sugar transporter
3	2245	55.1	729	2 T45780	sugar transporter-
4	2183	53.6	734	2 T51139	sugar transport pr
5	587.5	14.4	457	2 E70070	metabolite transpo
6	536	13.2	580	2 D86426	hypothetical prote
7	533	13.1	461	2 D70073	metabolite transpo
8	524.5	12.9	580	2 D84772	probable sugar tra
9	520.5	12.8	582	2 F71431	hypothetical prote
10	518.5	12.7	469	2 H97064	probable sugar-pro
11	505	12.4	473	2 G69789	sugar transporter
12	493.5	12.1	508	2 G84564	probable sugar tra
13	493	12.1	493	2 A85433	sugar transporter
14	489	12.0	521	2 G84864	probable membrane
15	488	12.0	549	2 T14606	probable sugar tra
16	486.5	11.9	482	2 B69803	metabolite transpo
17	474	11.6	547	2 C84593	probable sugar tra
18	473.5	11.6	464	2 F69587	L-arabinose transp
19	465	11.4	419	2 E69888	metabolite transpo
20	462	11.3	612	2 B40538	myo-inositol trans
21	460.5	11.3	511	2 A84537	probable sugar tra
22	459.5	11.3	511	2 H84536	probable sugar tra
23	458	11.2	560	2 T51485	sugar transporter-
24	451	11.1	464	2 F65079	galactose-proton s
25	448	11.0	464	2 C91106	galactose-proton s
26	448	11.0	464	2 F85951	galactose-proton s
27	445	10.9	464	2 AC0877	galactose-proton s
28	435.5	10.7	471	2 AB0868	L-arabinose isomer
29	434.5	10.7	472	2 S47089	arabinose-proton s

30	434	10.7	491	2 B86096	xylose-proton symp
31	434	10.7	491	2 F91255	xylose-proton symp
32	434	10.7	491	2 A26430	xylose transport p
33	432.5	10.6	490	2 T14545	probable sugar tra
34	432	10.6	557	2 T38125	myo-inositol trans
35	431	10.6	547	2 A48442	membrane transport
36	429.5	10.5	472	2 B91091	L-arabinose isomer
37	429.5	10.5	472	2 E85936	L-arabinose isomer
38	429.5	10.5	472	2 B26430	L-arabinose isomer
39	421.5	10.3	584	2 S69555	myo-inositol trans
40	418.5	10.3	468	2 S10014	glucose transport
41	417.5	10.2	555	2 S69671	hypothetical prote
42	415	10.2	606	2 T27072	hypothetical prote
43	412	10.1	487	2 E96782	hypothetical prote
44	412	10.1	502	2 B70845	probable sygar tra
45	407.5	10.0	613	2 T27077	hypothetical prote

ALIGNMENTS

RESULT 1

T06127
probable sugar transport protein F23E12.140 - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 09-Jul-2004
C/Accession: T06127
R/Bevan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Hoheisel, submitted to the Protein Sequence Database, April 1999
A/Reference number: Z15485
A/Accession: T06127
A/Molecule type: DNA
A/Residues: 1-729 <BEV>
A/Cross-references: UNIPROT:O65497; EMBL:AL022604; GSPDB:GN00062; ATSP:F23E12.140
A/Experimental source: cultivar Columbia; BAC clone F23E12
C/Genetics:
A/Gene: ATSP:F23E12.140
A/Map position: 4
A/Introns: 27/1; 216/1; 433/1; 519/3; 582/2
C/Keywords: sugar transport; transport protein

Query Match 63.4%; Score 2584; DB 2; Length 729;
Best Local Similarity 68.8%; Pred. No. 3e-172;
Matches 512; Conservative 89; Mismatches 123; Indels 20; Gaps 8;

Qy	56	MSGAVLVAIVASIGNLQGDNATIAAAVLYIKKFPQLQNEPTVEGLIVSMSLIGATIVT	115
Db	1	MSGAVLVAIAAAGVNLQGDNATIAAGAVLYIKKFNLESNPESVEGLIVAMSLIGATLIT	60
Qy	116	TFSGPLSDSIGRRPMLILSSILYFFSGLIMLWSPNVYVLLARFVDGFGIGLAVTLVPLY	175
Db	61	TCSGGVADWLGRRPMLILSSILYFVGLVLMWSPNVYVLLGRLLDGFVGLVTLVPIY	120
Qy	176	ISEIAPSEIRGLLNTLPQFSGSGGMFLSYCMVFGMSLSPSPDWRLMGLVLAIPSLFFFL	235
Db	121	ISEIAPSEIRGLLNTLPQFSGSGGMFLSYCMVFGMSLSPSPDWRLMGLVLAIPSLFFFL	180
Qy	236	TIYLPESPRWLVSCKGRMAEAKKVLQKLRGKDDVSGELSLLEGLVGGDTSEIYIIGP	295
Db	181	TVFPLPESPRWLVSCKGRMLEAKRVLQRLRGREDVSGEMALLVEGLIGGETTIEYIIGP	240
Qy	296	ATEAADDLVTDGDKKEQITLYCPPEEGQSWIARPSKPIMLGSLVSLASRHGS-MVNQSVPL	354
Db	241	ADEVTDHDIADVCKQIKLYCAEEGLSWARPVKG----GSTMSVLSRHGSTMRRQGS	296
Qy	355	MDPIVTLFGSVHENMPQAGGSMRSTLFPNFGSMFSTVDQHAQNEQWDEENLHRDDEYAS	414
Db	297	IDPLVTLFGSVHEKMPDT-GSMRSALFPFHFGSMFSGVGNQQRHEDWDENLVGEGEDYPS	355
Qy	415	DGAGGDEYEDNLHSPLLSRQATGAEGKDIVHHGRGSGALSMMRQTLLEGGDGVSSDTIGG	474
Db	356	D-HGDDSEDDLHSPILISRQTSME-KDMPHTAHGTLSTFRHGSQVQGAQGEAGSMGIGG	413

QY 475 GWLAWKSEKEGNGRKEGKFRVYLHQEGVPGSGRRGSIIVSLPGGGDVFESEFVHAAA 534
Db 414 GWQVAKWTEREDESQKE-----EGFPGSGRRGSIIVSLPGGGDTGE-ADFVQASA 462
QY 535 LVSQSALFSKGLAEPRMSDAAAMVHPSEVAAGKSRWKDLFEPCVRRALLVGVGIQILQQFA 594
Db 463 LVSQPALYSKDLKKEHTIGPAMVHPSE-TTKGSIWHDLHPGVKRALVGVGLQILQQFS 521
QY 595 GINGVLYTPQILEQAGVAVILSKFGLSSASASILISSLTLLMLPCIGFAMLMDLSGR 654
Db 522 GINGVLYTPQILEQAGVAVILSKFGLSSASASILISSLTLLMLPCIGFAMLMDLSGR 581
QY 655 RFLLLGTHPIILIASLVILVSNLIDGLTLAHALLSTVSVIVYFCCFVMGFGPIPNILCAE 714
Db 582 RTLLLTTPILIASLLVLVISNLVHMNSIVHAVLSTVSVVLYFCCFVMGFGPAPNILCSE 641
QY 715 IFPTRVRGLCIAICAFTEFWIGDIIVTYSPLVNLNAIGLAGVFSIYAVVCLISFVFLKV 774
Db 642 IFPTRVRGLCIAICALTFWICDIIVTYSPLVNLNAIGLAGVFSIYAVVCLISFVFLKV 701
QY 775 PETKGMPLVITEFFFAVGAKQAAA 798
Db 702 PETKGMPLVITEFFSVGARQAEA 725
RESULT 2
H86340
Sugar transporter [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: H86340
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: H86340
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-734 <STO>
A;Cross-references: UNIPROT:Q9SYQ3; GB:AE005172; NID:g4836905; PIDN:AAD30608.1; GSPDB:GN
C;Genetics:
A;Map position: 1

Query Match 55.9%; Score 2278; DB 2; Length 734;
Best Local Similarity 61.2%; Pred. No. 6.7e-151;
Matches 462; Conservative 112; Mismatches 145; Indels 36; Gaps 14;
QY 56 MSGAVLVAIVASIGNLLQGWDNATIAAAVLYIKKEFQIQNEPT-VEGLIVSNLSIGATIV 114
Db 1 MKGATLVALAATIGNFLQGWDNATAGAMVYINKDLNL---PTSVQGLVVMVSLIGATVI 57
QY 115 TTFSGPLSDSIGRRPMLILSSILYFFSGLIMLWSPNVYVLLARFVDGFGIGLAVTLVPL 174
Db 58 TTCGPISDWLGRRRPMLILSSVYVFCGLIMLWSPNVYVLCFARLLNGFAGLAVTLVPV 117
QY 175 YISEAPSEIRGLLNTLPQFSGGGMFLSYCMVFMGMSLSPSPDWIRMLGVLAIPSLFFFG 234
Db 118 YISEAPSEIRGLLNTLPQFSGGGMFLSYCMVFMGMSLSPSPDWIRMLGVLAIPSLFFFG 177
QY 235 LTIFVLPESPRWLVSCKGRVAAKVLQKLRGKDDVSGELSLLEGLVGGDTSIIEYIIG 294
Db 178 LTIFVLPESPRWLVSCKGRVAAKVLQKLRGKDDVSGELSLLEGLVGGDTSIIEYIIG 237
QY 295 PATEAADLVTGDKE-QITLYCPEEGQSWIARPSKGPIMLGSVLSLASHGSMVNQSV 353
Db 238 LEDHEGDDTLETVDDEGQRLYGTHTENQSYLARVPVPEQ---NSSLGLSRHSGSLANQSMI 294

QY 354 LMDPIVTLFGSVHENMPQAGGSMRSTLFPNFGSMFSTVDQ--HAKNEQWD---EENLHRD 408
Db 295 LKDLPLVNLFGSLHEKMPAEGGNTSGIFPHFGSMFSTADAPHGKPAHWEKDIESHYNKD 354
QY 409 DEEYASDGAGGDYED---NLHSPLLSRQATGAEGKDIVHGHGRGSALSMMRRTLLGEGGD 465
Db 355 NDDYATDDGAGDDDDSDNDRSLMSRQTTSDM-KDMLPHPTSGSTLSMRHSTLMQ-GN 412
QY 466 GVSSTDIQGGWQLAWKSEKEGNGRKEGKFRVYLHQEGVPGSGRRGSIIVSLPGGGDVFE 525
Db 413 GESSMGIGGGWHMGRYENDE-----YKRYLKEGDGAE-SRRGSIISIPGGPD--G 460
QY 526 GSEFVHAAALVSQSALFSKGLAEPRMSDAAAMVHPSEVAAGKSRWKDLFEPCVRRALLVGV 585
Db 461 GGSYTHASALVSRSVLGPKS-----VHGSAMVPEKIAASGPLMSALLEPGVVRALVVGV 515
QY 586 GIQILOQFAGINGVLYTPQILEQAGVAVILSKFGLSSASASILISSLTLLMLPCIGFA 645
Db 516 GIQILOQFSGINGVLYTPQILERAGVDILLSSGLSSISASFLISGLTTLMLPAIVVA 575
QY 646 MLLMDLSGRRFLLGTPIILIASLVILVSNLIDGLTLAHALLSTVSVIVYFCCFVMGFG 705
Db 576 MRLMDVSGRRSLLLTIPVLIVSLVVLVISELIHISKVNVNAALSTGCVVLYFCFFVMGYG 635
QY 706 PIPNILCAEIFPTRVRGLCIAICAFTEFWIGDIIVTYSPLVNLNAIGLAGVFSIYAVVCLI 765
Db 636 PIPNILCSEIFPTRVRGLCIAICAMVFWIGDIIVTYSPLVNLSSIGLVGVFSIYAAVCVI 695
QY 766 SFVVFVLKVPETKGMPLVITEFFFAVGAKQAAA 800
Db 696 SWIFVVMKVPETKGMPLVITEFFFAFGA-QAQA 729
RESULT 3
T45780
sugar transporter-like protein - Arabidopsis thaliana
N;Alternate names: protein F26013.130
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
C;Accession: T45780
R;Delseny, M.; Berger, C.; Cooke, R.; Grellert, F.; Laudie, M.; Mewes, H.W.; Lemcke, K.;
submitted to the Protein Sequence Database, December 1999
A;Reference number: Z23013
A;Accession: T45780
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-729
A;Cross-references: UNIPROT:Q9SD00; EMBL:AL133452
A;Experimental source: cultivar Columbia; BAC clone F26013
C;Genetics:
A;Map position: 3
A;Introns: 27/1; 216/1; 523/3; 578/2
A;Note: F26013.130

Query Match 55.1%; Score 2245; DB 2; Length 729;
Best Local Similarity 60.4%; Pred. No. 1.3e-148;
Matches 457; Conservative 110; Mismatches 144; Indels 46; Gaps 14;
QY 56 MSGAVLVAIVASIGNLLQGWDNATIAAAVLYIKKEFQIQNEPTVEGLIVSNLSIGATIV 115
Db 1 MRSVVLVALAAAIGNMLQGWDNATAGAVIYIKKEFHEKEPKIEGLIVAMSLIGATLIT 60
QY 116 TTFSGPLSDSIGRRPMLILSSILYFFSGLIMLWSPNVYVLLARFVDGFGIGLAVTLVPL 175
Db 61 TTFSGPVDKVGRRSRLILSSVLYFLSSIVFMFWSNVYVLLFARLLDGFGLAVTLVPIY 120
QY 176 ISETAPSEIRGLLNTLPQFSGGGMFLSYCMVFMGMSLSPSPDWIRMLGVLAIPSLFFFG 235
Db 121 ISETAPSEIRGLLNTLPQFSGGGMFLSYCLVFGMSLQSPSWRLMLGVLSIPSIAYFVL 180
QY 236 TIFVLPESPRWLVSCKGRVAAKVLQKLRGKDDVSGELSLLEGLVGGDTSIIEYIIG 295

Db 181 AAFPLPESPRWLVS KGRMDEARQVLQRLGRREDVSGELALLVEGLGVGKDTSEIEYVICP 240

QY 296 ATEAADDLVTDGKEQITLYGPEEGQSWIARPSKGPIMLSGLSVLSASRHGSMVNQSVPLM 355

Db 241 DNEENEGGNELPRKDOI KLYGPEDEGQSWNAKPVKGQ-----SSLALASRQGSMLPRGGSIM 296

QY 356 DPIVTLFGSVHENMP--QAGGSMRSTLFPNFGSMFSVTDQHAKNQEWDEENLHRDDEEVA 413

Db 297 DPLVTLFGSIHENLPSENMNASSRSMLPFNMGSI LGMGR--QESQWDPE---RNNED-- 349

QY 414 SDGAGGDYEDNLHSPLLSRQATGAEGKDIVHHGHRGSALSMMRQTL-LGEGGDGVSSTDI 472

Db 350 ----SSDQDENLNSPLLSPTT--EPDD--YHQRVTGTMHRRQSSLFMANVGETATATSI 401

QY 473 GGGWQLAWKWSKEGGRK-EGGFKRVYLHQE-----GVPGRRGSIIVSL-PGGGDV 523

Db 402 GGGWQLAWKYNDKVGADKRVNGGLQRMYIHEETANNNTNNIPFSRRGSLLSFHPEGDGH 461

QY 524 FEGSEFVHAAALVSQSALFSKGLAEPRMSDAAMVHPSEVAAKGSRWKDLFEPGVRRALLV 583

Db 462 DQVNGYVQAAALVSQASMPGKGGETAML-----PKEV-KDGPQWRELKEPQVKRALMV 514

QY 584 GVGILQIQFAGINGVLYYTPQILEQAGVAVILSKFGLSSASASILISLTLTLLMLPCIG 643

Db 515 GVGILQIQFAGINGVWYYTPQILEETGVSSLLTNLGISAESASLISALTLLMLPCI- 573

QY 644 FAMLLMDLSGRRFLLLGTIPILIASLVILVSVNLIDLGLTAHALISTVSVIVYFCCFVMG 703

Db 574 -----LVSMRSLMSTIPILILSLVTLVIGSLVNLGGSINALISTASVTVYLSCFVMG 626

QY 704 FGPIPNILCAEIPFTRVRGLCIACAFTFWIGDIIIVTYSLPVMLNAIGLAGVFSIYAVVC 763

Db 627 FGAIPNILCSEIFPTSVRGLCITICALTFWICDIIIVTITLPMKLSIGIAGVFGIYAIVC 686

QY 764 LISFVFLKVPETKGMPLVITEFFFAVGAKQAAKA 800

Db 687 AVAWFVFLKVPETKGMPLVISEFFFSVGAKQQDAAA 723

RESULT 4

T51139

sugar transport protein [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 28-Jul-2000 #sequence_revision 28-Jul-2000 #text_change 09-Jul-2004

C;Accession: T51139

R;Tjaden, J.; Neuhaus, E.

submitted to the EMBL Data Library, August 1995

A;Description: A new sugar transport protein from Arabidopsis thaliana.

A;Reference number: Z25311

A;Accession: T51139

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-734 <TJA>

A;Cross-references: UNIPROT:Q96290; EMBL:Z50752; PIDN:CAA90628.1

Query Match 53.6%; Score 2183; DB 2; Length 734;

Best Local Similarity 59.5%; Pred. No. 2.9e-144;

Matches 449; Conservative 110; Mismatches 160; Indels 36; Gaps 14;

QY 56 MSGAVLVAIVASIGNLLQGDNDATIAAAVLVIKKEFQLQNEPT-VEGLIVMSLSIGATIV 114

Db 1 MKGATLVALAATIGNFLQGDNDATIAGAMVYINKDLNL---PTSVQGLVAVMSLIGATIVI 57

QY 115 TTFSGPLSDSIGRRPMLILSSILYFFSGLINLWSPNVYVLLLARFVDFGFIGLAVTLVPL 174

Db 58 TTCSGPISDWLGRRPMLILSSVMYFVCGLINLWSPNVYVLCFARLNGFGAGLAVTLVPV 117

QY 175 YISEIAPSEIRGLNLTLPQFSGSGMFSLSCYMWFGMSLSPSPDWIRIMLGVLAIPSLFFFG 234

Db 118 YISETAPPEIRGQNLTLPQFLGSGGMFLSYCMVFTMSLSDSPSWRAMLVLSIPSLLYLF 177

QY 235 LTIFYLPESPRWLVS KGRMAEAKVKLQKLGKDDVSGELSLLLLEGLEVGGDTSEIEYIIG 294

Db 178 LTVFYLPESPRWLVS KGRMDEAKRVLQQLCGREDVTDDEALLVEGLDIGEKTMEDLLVT 237

QY 295 PATEAADDLVTDGKE-QITLYGPEEGQSWIARPSKGPIMLSGLSVLSASRHGSMVNQSVP 353

Db 238 LEDHEGDDTLFTVDEDDGQIRLYGTHENQSIARPVPEQ---NSSLGLSRSHGSLANQSMI 294

QY 354 LMDPIVTLFGSVHENMPQAGGSMRSTLFPNFGSMFSVTDQ--HAKNEQWD---EENLHRD 408

Db 295 LKDPLVNLFGSLHEKMPACGNTRSGIFPHFGSMFSTTADAPHGKPAHWEKDIESHYNKD 354

QY 409 DEEYASDGAGGDYED---NLHSPLLSRQATGAEGKDIVHHGHRGSALSMMRQTLGEGGD 465

Db 355 NDDYATDDGAGDDDDSDNDLSRPLMSRQITSM-D-KDMI PHPTSGSTLSMRHSTLMQ-GN 412

QY 466 GVSSTDIGGGWQLAWKWSKEGGRKGGFKRVYLHQEGVPGSRRGSIIVSLPGGGDVFE 525

Db 413 GESSMIGGGWHMGYRYENDE-----YKRYYLKEDGAE-SRRGSIISIPGGPD--G 460

QY 526 GSEFVHAAALVSQSALFSKGLAEPRMSDAAMVHPSEVAAKGSRWKDLFEPGVRRALLVG 585

Db 461 GGSYIHASALVSRSVLGPKS-----VHGSAMVPPEKIAASGPLWSALLEPGVKRALVVGV 515

QY 586 GIQILQIQFAGINGVLYYTPQILEQAGVAVILSKFGLSSASASILISLTLTLLMLPCIGFA 645

Db 516 GIQILQIQFSGINGVLYYTPQILERAGVDILSSLSGLSSI SASFLISGLTLLMLPAIVVA 575

QY 646 MLLMDLSGRRFLLLGTIPILIASLVILVSVNLIDLGLTAHALISTVSVIVYFCCFVMGFG 705

Db 576 MRLMDVSGRRSLILWTIPVLIVSLVVLVISELTHISKVNAALSTGCVVLYFCFFVMGYG 635

QY 706 PIPNILCAEIPFTRVRGLCIACAFTFWIGDIIIVTYSLPVMLNAIGLAGVFSIYAVVCLI 765

Db 636 PFQTSSVLKSSQQADRGLCIAICAMVFWIGDIIIVTYSLPVLLSSIELVGVPFSIYAAVCVI 695

QY 766 SFVFLKVPETKGMPLVITEFFFAVGAKQAAKA 800

Db 696 SWIFVYMKVPETKGMPLVITDYFAFGA-QAQASA 729

RESULT 5

E70070

metabolite transport protein homolog ywtG - Bacillus subtilis

C;Species: Bacillus subtilis

C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004

C;Accession: E70070

R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertero, C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Choi, A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.

Nature 390, 249-256, 1997

A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallercia, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.; Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A.; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mausel, Y., M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon, A.; Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror, akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.

A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.

A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

A;Reference number: A69580; MUID:98044033; PMID:9384377

A;Accession: E70070

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-457 <KUN>

A;Cross-references: UNIPROT:P96742; GB:Z99122; GB:AL009126; NID:G2636029; PIDN:CAB15600.1

A;Experimental source: strain 168

C;Genetics:

A;Gene: ywtG

C;Superfamily: glucose transport protein

Query Match 14.4%; Score 587.5; DB 2; Length 457;

Best Local Similarity 22.8%; Pred. No. 3e-33;

Matches 168; Conservative 102; Mismatches 162; Indels 305; Gaps 12;
QY 55 KMSGAVLVAIVASIGNLLQGWDNATIAAAVLYIKKEFQIQNEPTVEGLIVSMISLIGATIV 114
Db 2 KKQSNWLYFFGALGGALYGYDTGVISGAILFMKKELGL--NAFTEGLVSSLLVGAIG 59
QY 115 TTFSGPLSDSIGRRPMLILSSILYFFSGGLIMLWSPNVVYLLARFVDGFGIGLAVTLVPL 174
Db 60 SGAAGKLTDRFRGKKAIMAAALFCIGGLGVALAPNTGVMLFRILGLAVGTTIVPL 119
QY 175 YISEIAPSEIRGLLNTLPQFSGGGMFLSYCMVFGMSLSPSPDWRIMLGVLAPSLFFFG 234
Db 120 YLSELAPKHKRGALSSNLQMLMTVGIILSY--IVNYIFADAEAWRWMLGLAAPSLLLL- 176
QY 235 LTIFYLPSPRWLVSKGRMAEAKVQLKRGKDDVSGELSLLEGLVGGDTSEIEYIIG 294
Db 177 IGIIFMPSPRWLFTNGEESKAKILEKLRGTDOI----- 211
QY 295 PATEAADDLVTDGKEQITLYGPEEGQSWIARPSKGPIMLGSVLSASRHGSMVNQSVPL 354
Db 212 ----- 211
QY 355 MDPIVTLFGSVHENMPQAGGSMRSTLFPNFGSMFVTDQHAKNEQWDEENLHRDDEEYAS 414
Db 212 -----DQE-- 214
QY 415 DGAGGYEDNLHSPLLSRQATGAEGKDIVHHGHRGSALSMMRQTLLEGGDVSSDTIGG 474
Db 215 -----IH-----DI-- 218
QY 475 GWQLAWKWKSEKEGENRKEGKFRVYLHQEGVPGSRGSIIVSLPGGGDVFESEFVHAAA 534
Db 219 -----KEAEKQDEGLKE----- 231
QY 535 LVQSALFSGGLAEPRMSDAAVHPSEVAAGSRWKDLFEPGVRRALLVGVGIQLQOFA 594
Db 232 -----LFDPPWRPALIAGLGLAFLOQFI 254
QY 595 GINGVLYYTPQILEQAGVAVILSKFGLSSASAIL-----ISSLTLLMLPCIGFAMLLMD 650
Db 255 GTNTIYYAPKTFTNVG-----FG--NSASILGVIGIGTVNVLMTL-----VAIKIID 300
QY 651 LSGRRFLLGTIPILIASLVLVVSNLIDLTGLAHALLSTVSVIVYFCCFVMGEGPIPI 710
Db 301 KIGRKPLLLFGNAGMVISLVLALVNLFFNTPAASWTVICLVGFIVVFAVSWGPPVWV 360
QY 711 LCAEIFPFRVRLGICIACAFTWIGDIIIVTYSLPVMLNAIGLAVFSIYAVVCLISFVFV 770
Db 361 MLPELPLHVRGIGTGVSTMLHVGTLLIVSLTYPILMEAGISYFLIYAAIGIMAFLEV 420
QY 771 FLKVPETKGMPLLEVITE 787
Db 421 RFKVTETKGRSLEIEQ 437

RESULT 6
D86426
hypothetical protein F12P21.2 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: D86426
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: D86426

A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-580 <STO>
A:Cross-references: UNIPROT:Q9C757; GB:AE005172; NID:g11120774; PIDN:AAG30955.1; GSPDB:GN
C:Genetics:
A:Map position: 1
C:Superfamily: glucose transport protein
Query Match 13.2%; Score 536; DB 2; Length 580;
Best Local Similarity 21.6%; Pred. No. 1.6e-29;
Matches 175; Conservative 93; Mismatches 169; Indels 372; Gaps 11;
QY 62 VAIVASIGNLLQGWDNATIAAAVLYIKKEFQIQNEPT-VEGLIVSMISLIGATIVTTFSGP 120
Db 31 LAFSAGIGGLLFGYDTGVISGALLYIRDDFKSVDRNTWLQEMIVSMVAVAGAIYGAAGW 90
QY 121 LSDSIGRRPMLILSSILYFFSGGLIMLWSPNVVYLLARFVDGFGIGLAVTLVPLVISEIA 180
Db 91 ANDKLGRRSAILMADFLFLGAILMAAAPNPSLLVVGRVVFVGLGVGMASMTAPLYISEAS 150
QY 181 PSEIRGLLNTLPQFSGGGMFLSYCMVFGMSLSPSPDWRIMLGVLAIPLSFFFGTIFYL 240
Db 151 PAKIRGALVSTNGFLITGGQFLSYLINLAFT-DVTGTRWMLGIAGIPALLQFVL-MFTL 208
QY 241 PESPRWLVSKGRMAEAKVQLKRGKDDVSGELSLLEGLVGGDTSEIEYIIGPATEAA 300
Db 209 PESPRWLVRKGRREEAKAILRRIYSAEDVEQEIRALKDSVET----- 250
QY 301 DDLVTDGDKQITLYGPEEGQSWIARPSKGPIMLGSVLSASRHGSMVNQSVPLMDPIVT 360
Db 251 -EILEEGSSEKINMI----- 264
QY 361 LFGSVHENMPQAGGSMRSTLFPNFGSMFVTDQHAKNEQWDEENLHRDDEEYASDGAGD 420
Db 265 ----- 264
QY 421 YEDNLHSPLLSRQATGAEGKDIVHHGHRGSALSMMRQTLLEGGDVSSSTDIGGQWLAW 480
Db 265 -----KLCKAKT----- 271
QY 481 KWSEKEGENRKEGKFRVYLHQEGVPGSRGSIIVSLPGGGDVFESEFVHAAALVSQSA 540
Db 272 ----- 271
QY 541 LFSKGLAEPRMSDAAVHPSEVAAGSRWKDLFEPGVRRALLVGVGIQLQOFAINGVL 600
Db 272 -----VVRGLIAGVGLQVQFVGINTVM 295
QY 601 YYTPQILEQAGVAVILSKFGLSSASAILSSLTLLMLPCIGFAMLLMDLSGRFLLLG 660
Db 296 YYSPTIVQLAGFA-----SNRTALLSLVLTAGLNAFGSIISYIFIDRIGRKKLLII 346
QY 661 TIPILIASLVIL-----VVSNLIDLT----- 682
Db 347 SLFGVILSLGILTVFYEATHAPAISSLETQRFNNISCPDYKSAMNTNAWDCMCLKAS 406
QY 683 -----LAH-----ALLSTVSIVIVY 696
Db 407 SPSCGYCSSPIGKEHPGACWISDDSVKDLCHNENRLWYTRGCPSPNFGWFALLGLGLYIIF 466
QY 697 FCCFVNGFGPIPNILCAEIPFTRVRLGICIACAFTWIGDIIIVTYSLPVMLNAIGLAVF 756
Db 467 FS---PGMGTVPVWVNSEIYPLFRFGICGGIAATANWISNLIVAQSFLSLTEAGTSWTF 523
QY 757 SIYAVVCLISFVFVFLKVPETKGMPLLEVI 785
Db 524 LIFGVISVIALLFVWVCVETKGMPEEI 552

RESULT 7
D70073
metabolite transport protein homolog yxuC - Bacillus subtilis
C:Species: Bacillus subtilis

C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C:Accession: D70073
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berten
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallen
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Togononi, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: D70073
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-461 <KUN>
A:Cross-references: UNIPROT:P46333; GB:299124; GB:AL009126; NID:g2636442; PIDN:CAB16017.
A:Experimental source: strain 168
C:Genetics:
A:Gene: yxcC
C:Superfamily: glucose transport protein

Query Match 13.1%; Score 533; DB 2; Length 461;
Best Local Similarity 21.6%; Pred. No. 1.9e-29;
Matches 157; Conservative 102; Mismatches 168; Indels 300; Gaps 12;
QY 61 LVAIVASIGNLLQGDNDATIAAAVLYIKKEFQLQNEPTVEGLIVSMISLIGATIVTTFSGP 120
DB 9 MIYFFGALGGLLYGDTGVISGALLFINNDIPLTT--LTEGLVVSMLLLGAIFGSALSCT 66
QY 121 LSDSIGRRPMLILSSILYFFSGGLMLWSPNVVLLARFVDFGFGIGLAVTLVPLYISEIA 180
DB 67 CSDRWGRKRVVFLSIIIFIIGALACAFSQTIGMLIASRVILGLAVGSTALVPVYLSEMA 126
QY 181 PSEIRGLNLTLPQFSGSGGMFLSYCMVFGMSLSPSPDWRIMLGVLAIPLSFFGLTIFYL 240
DB 127 PTKIRGTGLTNNLMIVTGILLAY--IVNLYFTPFPAWRWTVGLAAVPAVLL--IGIAFM 183
QY 241 PESPRWLVSCKRMAEAKKVLQKRGKDDVSGELSLLEGLVGGDTSEIEYIIGPATEAA 300
DB 184 PESPRWLVSCKRMAEAKKVLQKRGKDDVSGELSLLEGLVGGDTSEIEYIIGPATEAA 222
QY 301 DDVTDGDKQITLYGPEEGQSWIARPSKGPIMLGSVLSASRHGSMVNSQSVPLMDPIVT 360
DB 223 -----EAEEKETTL----- 231
QY 361 LFGSVHENMPQAGGSMRSTLFPNFGSMFVTDQHAKEQWDEENLHRDDEEYASDGAGGD 420
DB 232 ----- 231
QY 421 YEDNLHSPLLSRQATGAEGKDIVHHGHRGSALSMRRQTLLGEGGDVSTDIGGWQLAW 480
DB 232 -----GV-----LKA 236
QY 481 KWSEKEGENRKEGGFKRVYLHQEGVPGSRGSIIVSLPGGGDVFESEFVHAAALVSQSA 540
DB 237 KW----- 238
QY 541 LFSKGLAEPRMSDAAVHPSEVAAGSRWKDLFEPGVRRALLVGVGIQLQFAGINGVL 600
DB 239 -----IRPMLLIGVGLAIFQQAAGINTVI 262
QY 601 YYTPQILEQAGVAVILSKFGLSSASASILISLTLLMLPCIGFAMLLMDLSGRRFLL-- 659
DB 263 YYAP-----TFTKAGLGTASALGMIGILNVIMCI--TAMILIDRVGRKKLLIW 312
QY 660 GTIPILIASLVILVSNLIDLGLTAH--ALLSTVSIVVYFCCFVMGFGIPNLCIEIFPT 718

DB 313 GSVGITLSLAALSGV--LLTLGLSASTAMTVVFLGVYIVFYQATGPPVWVLMPELPPS 370
QY 719 RVRGLCIACAFTWIGDIIIVTYSPLVMLNAIGLAGVFSIYAVVCLISFVFLKVPETK 778
DB 371 KARGAATGTFTLVLSAANLIVSLVPLMLSAMGIAWFMVFSVICLLSFFFAFYMVPEK 430
QY 779 GMPLEVI 785
DB 431 GKSLEEI 437
RESULT 8
D84772
probable sugar transporter [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: D84772
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; N
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: D84772
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-580 <STO>
A:Cross-references: UNIPROT:Q9ZQP6; GB:AE002093; NID:g4263781; PIDN:AAD15441.1; GSPDB:GN
C:Genetics:
A:Gene: At2g35740
A:Map position: 2
C:Superfamily: glucose transport protein

Query Match 12.9%; Score 524.5; DB 2; Length 580;
Best Local Similarity 20.9%; Pred. No. 1e-28;
Matches 171; Conservative 102; Mismatches 175; Indels 369; Gaps 13;
QY 62 VAIVASIGNLLQGDNDATIAAAVLYIKKEF--QLQNEPTVEGLIVSMISLIGATIVTTFSGP 120
DB 29 LALSAGIGGLLFGYNTGVIAGALLYIKKEFGEVDNKTWLQEIIVSMTVAGAAIGGW 88
QY 121 LSDSIGRRPMLILSSILYFFSGGLMLWSPNVVLLARFVDFGFGIGLAVTLVPLYISEIA 180
DB 89 YNDKFGRRMSVLIADVLFLGLAVMVAIAHAPVILGRLLVGVFGVGMASMTSPLYISEMS 148
QY 181 PSEIRGLNLTLPQFSGSGGMFLSYCMVFGMSLSPSPDWRIMLGVLAIPLSFFGLTIFYL 240
DB 149 PARIRGALVSTNGLLTIGGQFLSYLINLAFVHTPG--TWRWMLGVSAPAIQFCL--MLTL 206
QY 241 PESPRWLVSCKRMAEAKKVLQKRGKDDVSGELSLLEGLVGGDTSEIEYIIGPATEAA 300
DB 207 PESPRWLVRNDRKAESRDILERIYPAEWVEAFLAALKESVR-----AETA 251
QY 301 DDVTDGDKQITLYGPEEGQSWIARPSKGPIMLGSVLSASRHGSMVNSQSVPLMDPIVT 360
DB 252 DE----- 253
QY 361 LFGSVHENMPQAGGSMRSTLFPNFGSMFVTDQHAKEQWDEENLHRDDEEYASDGAGGD 420
DB 254 ----- 253
QY 421 YEDNLHSPLLSRQATGAEGKDIVHHGHRGSALSMRRQTLLGEGGDVSTDIGGWQLAW 480
DB 254 -----DIIGH----- 258
QY 481 KWSEKEGENRKEGGFKRVYLHQEGVPGSRGSIIVSLPGGGDVFESEFVHAAALVSQSA 540
DB 259 TFSDK-----LRGA----- 267
QY 541 LFSKGLAEPRMSDAAVHPSEVAAGSRWKDLFEPGVRRALLVGVGIQLQFAGINGVL 600
DB 268 -----LSNPVVRHGLAAGITVQAAQFVGINTVM 296

Db 193 -ITLFFVPESPRFLVKSNGIKKAAAVLTKINGAEIAKQELDSISKSLATENDSSL----- 246
QY 294 GPATEAADLVTGDKEQITLYPPEGQSWIARPSKGPIMLGSVLSLASRHGSMVNSQVSP 353
Db 247 -----GQ----- 248
QY 354 LMDPIVTLFGSVHENMPQAGSMRSTLFPNFGSMFSTVDQHAKEQWDEENLHRDDEYA 413
Db 249 ----- 248
QY 414 SDGAGGDYEDNLHSPLLSRQATCAEGKDIVHHGHRGSALSMMRQTLLGEGGDGVSSTDIG 473
Db 249 ----- 248
QY 474 GGWQLAWKWKSEKEGENGRKEGFKRVYLHQEGVPGSRRSISVLPGGGDVFEGSEFFVHAA 533
Db 249 ----- 248
QY 534 ALVQSALFSKGLAEPRMSDAMVHPSEVAAGSRWKDLFEPGVRRALLVGVGIGIQF 593
Db 249 -----LLQPLRRALLIGIFLAIFNQ 270
QY 594 AGINGVLYYTPQILEQAGVAVILSKFGLSSASASIL--ISSLTLLMLPCIGFAMLLMDL 651
Db 271 IGMNSITYGPEIFQMIGF-----KNSSFLATSVIGVVEFSTIL-----AMFLIDK 318
QY 652 SGRFFLL-LGTIPILIASLVILVNSLIDLGTLAHALLSTVSVI-VYFCFVMGF-----G 705
Db 319 LGRKKLMBIGSAAMAVFMLLI-----GTSFYIKLSNGFVILFIICFVVSFCISMG 369
QY 706 PIPNLCABIFPTRVRGLCIACAFTFWIGDIIIVTYSLPVMLNAIGLAGVFSIYAVVCLI 765
Db 370 PIPWIMPEIFPNHLRARATGIATIFLWGANWAIGQFTPMLLNGIGGAYTFWIFCGINVI 429
QY 766 SFVFEVLKVPETKGMPLVITEFFAVGAKQAA 797
Db 430 CFLVVTXVTPETKNKSLEEIEKFWIPKSKQNA 461

RESULT 11
G69789
sugar transporter homolog ydjK - Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C;Accession: G69789
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: G69789
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-473 <KUN>
A;Cross-references: UNIPROT:O34718; GB:Z99107; GB:AL009126; NID:g2632866; PIDN:CAB12442.
A;Experimental source: strain 168
C;Genetics:
A;Gene: ydjK
C;Superfamily: glucose transport protein

Query Match 12.4%; Score 505; DB 2; Length 473;

Best Local Similarity 22.3%; Pred. No. 1.8e-27;
Matches 168; Conservative 106; Mismatches 176; Indels 302; Gaps 16;
QY 54 DKMSGAVLVAIVASIGNLLQWMDNATIAAAVLYIKKEFQLONEPTVEGLIVMSLIGATI 113
Db 6 NQMSFLRTILVSTFGLLFGYDTGVNGALPYMGEPDQNLNAPTEGLVTSLLFGAAL 65
QY 114 VTTFSGLSDSIGRRPMLILSSILYFFSGLIMLWSPNVYVLLARFVDFGIGLAVTLVP 173
Db 66 GAVFGGRMSDFNGRRKNILFLAVIFISTIGCTFAPNVTVMISRFVLGIAVGASVTVP 125
QY 174 LYISEIAPSEIRGLLNTLPQFSGSGMFLSYCM--VFGMSLSP-SPDWRIMLVLAIPSL 230
Db 126 AYLAEMSPVESRGRMVTQNELMIVSQQLLAFVFNAILGTTMGDNSHVRFMLVIASLPAL 185
QY 231 F-FFGLTIFYLPESPRWLVSGRMAKAKVQLKRGKODVSGELSLLEGLVGGDTSIE 289
Db 186 FLFFGM--IRMPESPRWLVSGRKEDALRVLKKIR----- 218
QY 290 EYIIGPATEAADLVTGDKEQITLYGPEEQSWIARPSKGPIMLGSVLSASRHGSMVN 349
Db 219 ----- 218
QY 350 QSVPLMDPIVTLFGSVHENMPQAGSMRSTLFPNFGSMFSTVDQHAKEQWDEENLHRDD 409
Db 219 -----D 219
QY 410 EYASDGAGGDYEDNLHSPLLSRQATCAEGKDIVHHGHRGSALSMMRQTLLGEGGDVSS 469
Db 220 EKRAA-----AEQEI----- 230
QY 470 TDIGGWQLAWKWKSEKEGENGRKEGFKRVYLHQEGVPGSRRSISVLPGGGDVFEGSEF 529
Db 231 -----EFAPKK----- 236
QY 530 VHAAALVSQSALFSKGLAEPRMSDAMVHPSEVAAGSRWKDLFEPGVRRALLVGVGIGI 589
Db 237 -----EDQLEKAT-----FKDLSVDPVRRIVFVIGLIGIAI 265
QY 590 LQQFAGINGVLYYTPQILEQAGVAVILSKFGLSSASASILISSLTLLMLPCIGFAMLLM 649
Db 266 VQQITGVNSIMYGTETILRNSG---FQTEAALIGNIANGVISVLAFTV-----GIWLL 315
QY 650 DLSGRR-FLLLGTIPILIASLVILVNSLIDLGTLAHALLSTVSVIVYFCCFVMG-FGPI 707
Db 316 GRVGRRRPMLMTGLIGTTTALLLIGIFSLVLE-GSPALPYV-VLSLTVTFLAFQQAISPV 373
QY 708 PNILCAEIFFTRVRGLCIACAFTFWIGDIIIVTYSLPVMLNAIGLAGVFSIYAVVCLISF 767
Db 374 TWLMLSEIFPLRLRGLMGVTVFCLMVNVNPAVSFTFPILLAAIGLSTFTFFIFVGLGICSV 433
QY 768 VFVFLKVPETKGMPLVITEFFAV---GAKQ 795
Db 434 LFVKRFLPETKGLSLEQLEENFRAYDHSAGAK 465

RESULT 12
G84564

probable sugar transporter [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: G84564
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; N
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: G84564
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-508 <STO>
A;Cross-references: UNIPROT:Q9ZNS0; GB:AE002093; NID:g4218010; PIDN:AAD12218.1; GSPDB:GN

C:Genetics:		A:Status: preliminary	
A:Gene: At2g18480		A:Molecule type: DNA	
A:Map position: 2		A:Residues: 1-493 <STO>	
C:Superfamily: glucose transport protein		A:Cross-references: UNIPROT:O23213; GB:NC_001268; NID:g7270615; PIDN:CAB80333.1; GSPDB:G	
Query Match		12.1%; Score 493.5; DB 2; Length 508;	
Best Local Similarity		22.7%; Pred. No. 1.2e-26;	
Matches 170; Conservative 99; Mismatches 174; Indels 305; Gaps 19;		Matches 177; Conservative 102; Mismatches 178; Indels 306; Gaps 23;	
Qy	63	AIVASIGNLLQGDNATIAAAVLYIKKEFQLQNEPTVE---	GLIVMSLIGATIVTTFSG 119
Db	26	AIVASIIISIIIFGYDTGVMGSAQIFIRDDLKI-NDTQIEVLGILNLALVGSILT---	AG 80
Qy	120	PLSDSIGRRPMLILSSILYFFSGLIMLWSPNYYVLLARFVDGFGIGLAVTLVPLYISEI 179	
Db	81	KTSDIVGRRYTIALSAVIFLVGSLMGYPNYPVLMVGRCIAGVGGFALMIAPVYSAEI 140	
Qy	180	APSEIRGLLNTLPQFSGSGMFL---SYCMVFGMSLSPSPDWRIMLVLAIPSLFF-FG 234	
Db	141	SSASHRGFLTSLPELCISLGLILGYVSNYC--FG-KLTLKLGWRLMLGIAAPPSLILAFG 197	
Qy	235	LTIFVLPESPRWLVSKGMAEAKKVLQKLRGKDDVSGELSLLLLEGLEVGSDTSIEEYIIG 294	
Db	198	IT--RMPESPRWLVMQGRLEEAKKIM-----	221
Qy	295	PATEAADLVTGDKEQITLYGPEEGQSWIARPSKGPIMLGSVLSLASHGSMVNQSVPL 354	
Db	222	-----	221
Qy	355	MDPIVTLFGSVHENMPQAGSMRSTLFPNFGSMFVSVDQAKNEQWDEENLHRDDEEYAS 414	
Db	222	-----VLVSNTEEEAEER-----	234
Qy	415	DGAGGDYEDNLHSPLLSRQATGAEGKDIVHHGHSALSMMRQTLLEGGDVSSDIDGG 474	
Db	235	-----FRDIL-----TAAE-----VDVTEIKEVGG 254	
Qy	475	GWQLAWKWEKEGEGNKRKGGFKRVYLHQEGVPGSRRGSIVSLPGGGDVFEGSEFVHAAA 534	
Db	255	G-----VKKKNHGK-----	263
Qy	535	LVSQSALFSKGLAEPRMSDAAMVHPSEVAAKGRWKDLF---EPGVRRALLVGVGIQILQ 591	
Db	264	-----SVWRELVIKPRPAVRLILIAAVGIHFFE 291	
Qy	592	QFAGINGVLYTTPQILEQAGVA---VILSKFGLSSASASILISLTLMLPCIGFAML 647	
Db	292	HATGIEAVVLYSPRIPKAGVSVKDKLLLATVGVGLTKAFFII-----IATF 338	
Qy	648	LMDLSGRRFLLLTGTPILIASLVILVSNLI--DLGTLAHL-LSTVSVIVYFCFVMGF 704	
Db	339	LLDKVGRKLLLTSTGGMVFALTSLAVSLTMVQRFGRLAWALSISIVTYAFVAFSISGL 398	
Qy	705	GPINILCAEIFFPTRVRGLCIAICAFTFWIGDIIVTYSLPVMLNAIGLAGVFSIYAVVCL 764	
Db	399	GPITWVYSSEIFFPLRLRAQAGASIGVAVNRIMNATVSMFSLMTKAITTGGVFFVFAGIAV 458	
Qy	765	ISFVVFVLKVPETKGMPLVITEFFAVG 792	
Db	459	AAWVFFFMPLPETKGLPLEMEKLFGGG 486	
RESULT 13		A85433	
sugar transporter like protein [imported] - Arabidopsis thaliana		C:Species: Arabidopsis thaliana (mouse-ear cress)	
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004		C:Accession: A85433	
R;anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring Nature 402, 769-777, 1999		A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.	
A:Reference number: A85001; MUID:20083488; PMID:10617198		A:Accession: A85433	
Query Match		12.1%; Score 493; DB 2; Length 493;	
Best Local Similarity		23.2%; Pred. No. 1.3e-26;	
Matches 177; Conservative 102; Mismatches 178; Indels 306; Gaps 23;		Matches 177; Conservative 102; Mismatches 178; Indels 306; Gaps 23;	
Qy	48	SDDILEDKMSG---AVLVAIVASIGNLLQGDNATIAAAVLYIKKEFQLQNEPTVE--- 100	
Db	2	ADQISGEKPAGVNRFAALQCAIVASIVSIIFGYDTGVMGAMVFIKEDLK-TNDVQIEVLT 60	
Qy	101	GLIVMSLIGATIVTTFSGPLSDSIGRRPMLILSSILYFFSGLIMLWSPNYYVLLARFV 160	
Db	61	GILNLALVGSLL---AGRTSDIIGRRYTIVLASILFMLGSILMGWGNYPVLLSGRCT 116	
Qy	161	DGFIGLAVTLVPLYISEIAPSEIRGLLNTLPQFSGSGMFLSYCMVFGMSLSPSP-DWR 219	
Db	117	AGLGVGFALMVAPVYSAEIATASHRGLLASLPHLCISIGILLGYIVNYFFSKLPMHIGWR 176	
Qy	220	IMLVLAIPSLFF-FGLTIFVLPESPRWLVSKGMAEAKKVLQKLRGKDDVSGELSLLE 278	
Db	177	LMGLIAAVPSLVLAFG--ILKMPESPRWLIMQGRLEKEGKEIL----- 216	
Qy	279	GLEVGSDTSIEEYIIGPATEAADLVTGDKEQITLYGPEEGQSWIARPSKGPIMLGSVL 338	
Db	217	-----ELVSN-----SPEAE----- 227	
Qy	339	SLASHGSMVNQSVPLMDPIVTLFGSVHENMPQAGSMRSTLFPNFGSMFVSVDQAKNE 398	
Db	228	-----LR----- 229	
Qy	399	QWDEENLHRDDEEYASDGAGGDYEDNLHSPLLSRQATGAECK--DIVHHGHSALSMSR 455	
Db	230	-----FQD-----IKAAAGIDPKCVDVV----- 248	
Qy	456	RQTLLEGGDGVSSDIDGGWQLAWKWEKEGEGNKRKGGFKRVYLHQEGVPGSRRGSIV 515	
Db	249	-----KMEGK-----THGEGV----- 260	
Qy	516	SLPGGGDVFEGSEFVHAAALVSQSALFSKGLAEPRMSDAAMVHPSEVAAKGRWKDLF-- 573	
Db	261	-----WKELILR 267	
Qy	574	EPGVRRALLVGVGIQILQFAGINGVLYTTPQILEQAGVAVILSKFGLSSASASILISS 632	
Db	268	PTPAVRRVLLTALGIHFFQHASGIEAVLLYGPRIFFKAGITT-KDKLFLVTIGVGIM--- 323	
Qy	633	LTLLMLPCIGFAMLLMDLSGRRFLLLTGTPILIASLVIL-----VVSNLIDLGTLAHAL 687	
Db	324	KTTFIFT-----ATLLDKVGRKLLLTSGVGMVIALTMLGFLGTVAQNA--GGKLAWAL 376	
Qy	688	-LSTVSVIVYFCFVMGFPIPNILCAEIFFPTRVRGLCIAICAFTFWIGDIIVTYSLPVM 746	
Db	377	VLSIVAAYSFVAFSISGLGPITWVYSSEVFLKRAQAGSLGAVAVNRVMNATVSMFSLSL 436	
Qy	747	LNAIGLAGVFSIYAVVCLISFVFLKVPETKGMPLVITEFF 789	
Db	437	TSAITTGAFFMFAGVAAVAWNFFFLFLPETKGSLEIEALF 479	
RESULT 14		G84864	
probable membrane transporter [imported] - Arabidopsis thaliana		C:Species: Arabidopsis thaliana (mouse-ear cress)	
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004		C:Accession: G84864	
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;			

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999

A;Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: G84864
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-521 <STO>
A;Cross-references: UNIPROT:O22848; GB:AE002093; NID:g2289003; PIDN:AAB64332.1; GSPDB:GN
C;Genetics:
A;Gene: At2g43330
A;Map position: 2
C;Superfamily: glucose transport protein

T14606
probable sugar transport protein 205 - beet
C:Species: Beta vulgaris (beet)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C:Accession: T14606; T14617
R:Chiou, T.J.; Bush, D.R.
Plant Physiol. 110, 511-520, 1996
A>Title: Molecular cloning, immunochemical localization to the vacuole, and expression in
A:Reference number: Z18131; MUID:96351183; PMID:8742332
A:Accession: T14606
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-549 <CHI>
A:Cross-references: UNIPROT:P93075; EMBL:U64902; NID:g1778092; PID:g1778093
A:Accession: T14617
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-541,'SVQV', <CH2>
A:Cross-references: EMBL:U64903; NID:g1778094; PID:g1778095
C:Genetics:
A>Note: BvcdNA-205; BvcdNA-397
C:Superfamily: glucose transport protein
C:Keywords: sugar transport

Db 395 WAIALCITM-VCAVVASFSIGLGPITWVYSSEVFPRLRAQGTSMGV-AVNRVVGVISI 452
QY 741 YSLPVMNLNAIGLAGVFSIYAVVCLISFVFLKVPETKGMPLLEVITEFF 789
Db 453 FFLP-LSHKITTGGAFFLEFGGIAIIAWFFLTFLPETRGRTLENMHLEF 500

Search completed: October 13, 2004, 11:39:09
Job time : 26.286 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 13, 2004, 11:40:19 ; Search time 6494.01 Seconds
(without alignments)
4489.025 Million cell updates/sec

Title: US-10-051-909-32
Perfect score: 4075
Sequence: 1 IRSGSWLAVQTPTDRLRR.....PLEVITEFFAVGAKQAAKA 800

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2 1/USPTO_spool.h/US10051909/runat 13102004 123336 19917/app_query.fasta_1.1678
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10051909 @CGN 1 1 6263 @runat 13102004 123336 19917 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST: *
1: gb_est1: *
2: gb_est2: *
3: gb_htc: *
4: gb_est3: *
5: gb_est4: *
6: gb_est5: *
7: gb_est6: *
8: gb_gss1: *
9: gb_gss2: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3027	74.3	2180	3 AY105508	AY105508 Zea mays
2	1399	34.3	911	8 CC423708	CC423708 PUH049TB
3	1382.5	33.9	845	8 BZ723844	BZ723844 PUCFI60TD
4	1279	31.4	795	8 BZ989776	BZ989776 PUGDY07TD
5	1268.5	31.1	1135	3 CNS0A3H4	BX828912 Arabidops
6	1168	28.7	833	6 CA065405	CA065405 SCACAD103
7	1154	28.3	733	6 CA146057	CA146057 SCVPR207
8	1142	28.0	696	6 CA075335	CA075335 SCJFAM106
9	1120	27.5	1082	7 CK208781	CK208781 FGAS02050

10	1101	27.0	829	5 BQ752737	BQ752737 WHE4118_E
11	1100	27.0	871	6 CA170109	CA170109 SCQGSB108
12	1086	26.7	1107	5 BX841312	BX841312 BX841312
13	1066.5	26.2	848	7 CF445664	CF445664 EST682009
14	1063.5	26.1	769	6 CA243478	CA243478 SCQFL407
15	1055	25.9	994	6 CA067439	CA067439 SCQGLAD106
16	1034	25.4	636	6 CF018228	CF018228 QBM9a03.X
17	1033.5	25.4	950	7 CK263963	CK263963 EST710041
18	1030	25.3	692	6 CA112602	CA112602 SCEQLB106
19	1022	25.1	631	6 CA252785	CA252785 SCBFLL114
20	1018	25.0	787	6 CF182124	CF182124 JG-2 Chin
21	1014.5	24.9	643	6 CA263039	CA263039 SCFRLB202
22	1014.5	24.9	858	6 CA207025	CA207025 SCEQSB1C0
23	1009	24.8	644	6 CA279697	CA279697 SCCFL800
24	1003	24.6	616	6 CA081000	CA081000 SCCCAM200
25	997	24.5	606	6 CA230633	CA230633 SCJFFL3C0
26	990.5	24.3	745	6 CB635485	CB635485 OSIIEB150
27	981	24.1	594	1 AI861088	AI861088 603011H02
28	980	24.0	620	6 CA167255	CA167255 SCCCSB100
29	966	23.7	669	6 CA073177	CA073177 SCEPAM105
30	957	23.5	817	8 BZ819234	BZ819234 PUGDY07TB
31	945	23.2	903	6 CF211894	CF211894 CGF100066
32	941.5	23.1	655	6 CA268224	CA268224 SCQSR305
33	939	23.0	701	6 CD879843	CD879843 AZO4.1061
34	937	23.0	827	6 CA207188	CA207188 SCEQSB1C0
35	935	22.9	801	7 CK201790	CK201790 FGAS01031
36	930	22.8	859	4 BM817360	BM817360 HC106B05
37	928	22.8	749	5 BU040051	BU040051 PP Lea000
38	924	22.7	688	6 CA502649	CA502649 WHE4338_A
39	921.5	22.6	700	5 BQ870435	BQ870435 QGD9A19.Y
40	921.5	22.6	703	6 CA502490	CA502490 WHE4048_B
41	921	22.6	687	6 CD893252	CD893252 G118.123E
42	916	22.5	847	6 CB981615	CB981615 CAB70005
43	914	22.4	736	2 BF108319	BF108319 EI22F22_N
44	909	22.3	572	4 BM325261	BM325261 PIC1_42_E
45	909	22.3	793	6 CA266132	CA266132 SCJLRT308

ALIGNMENTS

RESULT 1
AY105508
LOCUS AY105508 2180 bp mRNA linear HTC 16-OCT-2002
DEFINITION Zea mays PC0114533 mRNA sequence.
ACCESSION AY105508
VERSION AY105508.1 GI:21208586
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 2180)
AUTHORS Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
TITLE Maize Mapping Project/Dupont Consensus Sequences for Design of
Overgo Probes
JOURNAL Unpublished (2002)
REFERENCE 2 (bases 1 to 2180)
AUTHORS Coe,E.H.
TITLE Direct Submission
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
COMMENT If you are interested in getting corresponding physical clones,
these are publicly available from ZmDB and may be found by BLAST
searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schnable, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.
LOCATION/Qualifiers
1. .2180
/organism="Zea mays"

FEATURES
source

/mol_type="mRNA"
/db_xref="MaizeDB:636667"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Cornsensus
Library"
/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"

ORIGIN

Alignment Scores:
Pred. No.: 1.02e-294 Length: 2180
Score: 3027.00 Matches: 591
Percent Similarity: 99.66% Conservative: 2
Best local Similarity: 99.33% Mismatches: 2
Query Match: 74.28% Indels: 0
DB: 3 Gaps: 0

US-10-051-909-32 (1-800) x AY105508 (1-2180)

QY 206 MetValPheGlyMetSerLeuSerProSerProAspTrpArgIleMetLeuGlyValLeu 225
DB 50 ATGGTGTGGGATGTCCTGCGCCATCACCGGATGAGAAATATGCTTGGTGTGCTC 109
QY 226 AlaileProSerLeuPhePheGlyLeuThrIlePheTyrLeuProGluSerProArg 245
DB 110 GCGATACCTTCATGTTCTTCTTGGTTTGACAAATATTTATCTTCTGAACTCCAAGA 169
QY 246 TrpLeuValSerLysGlyArgMetAlaGluAlaLysLysValLeuGlnLysLeuArgGly 265
DB 170 TGGCTCGTAGCAAGGTGGGATGGCAGAGGCAAAAGAGGTGTGCAAAAGTTACGGTGG 229
QY 266 LysAspValSerGlyGluLeuSerLeuLeuGluGlyLeuGluValGlyGlyAsp 285
DB 230 AAAGACGATGTCAGGTGAATGTCCTTCTTCTCGAAGGTTGGAGGTTGGAGGAGAC 289
QY 286 ThrSerIleGluGluTyrIleIleGlyProAlaThrGluAlaAlaAspAspLeuValThr 305
DB 290 ACTTCCATTGAAGAGTACATCATTTGGACCTGCCACCGAGGCGGATGATCTTGTACT 349
QY 306 AspGlyAspLysGluGlnIleThrLeuTyrGlyProGluGluGlyGlnSerTrpIleAla 325
DB 350 GACGGTGATAAGGAACAAATCACACTTATGGCCCTGAAGAAGGCCAGTCATGGATTGCT 409
QY 326 ArgProSerLysGlyProIleMetLeuGlySerValIleuSerLeuAlaSerArgHisGly 345
DB 410 CGACCTTCTAAGGACCCCATCATGCTTGAAGTGCTGCTTCTTGTGATCTCGTCATGGG 469
QY 346 SerMetValAsnGlnSerValProLeuMetAspProIleValThrLeuPheGlySerVal 365
DB 470 AGCATGGTGAAACCAGAGTGATCCCTTATGGATCCGATGTGACACTTTTGGTAGTGTG 529
QY 366 HisGluAsnMetProGlnAlaGlyGlySerMetArgSerThrLeuPheProAsnPheGly 385
DB 530 CATGAGAAATATGCCCTCAAGCTGGAGGAAGTATGAGGAGCACATTGTTTCCAAACTTTGGA 589
QY 386 SerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTrpAspGluGluAsnLeu 405
DB 590 AGTATGTCAGTGTACAGATCAGCATGCGCATGCCAAATGACAGTGGGATGAGAGAAATCTT 649
QY 406 HisArgAspAspGluGluTyrAlaSerAspGlyAlaGlyGlyAspTyrGluAspAsnLeu 425
DB 650 CATAGGATGACGAGGAGTACCGCATCTGATGGTGCAGGAGGTGACTATGAGGACAAATCTC 709
QY 426 HisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIleValHisHis 445
DB 710 CATAGCCCATTTGCTGTCCAGGAGGCAACAGGTGCGGAAGGAAGGACATTGTGCACCAT 769
QY 446 GlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeuGlyGlyAsp 465

DB 770 GGTCAACCGTGAAGTGTCTTTGAGCATGAGAAGGCAAGCCCTCTTAGGGAGGGTGGAGAT 829
QY 466 GlyValSerSerThrAspIleGlyGlyGlyTyrGlnLeuAlaTrpLysTrpSerGluLys 485
DB 830 GGTGTGACGACACTGATATCGTGGGGGATGGCAGCTTGCTTGGAAATGGTCAGAGAAG 889
QY 486 GluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGly 505
DB 890 GAAGGTGAGAAATGGTAGAAAGAGGAGGTGTTTCAAAAGAGTCTACTTGCACCAAGAGGGA 949
QY 506 ValProGlySerArgArgGlySerIleValSerLeuProGlyGlyGlyAspValPheGlu 525
DB 950 GTTCTGCTCAAGAAAGGGGCTCAATTGTTTCACTTCCCGGTGGTGGCGATGTTCTTGAG 1009
QY 526 GlySerGluPheValHisAlaAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGly 545
DB 1010 GGTAGTGAGTTTGTACATGCTGCTGCTTTAGTAAGTCAGTCAGCAGCACTTTTCTCAAAGGT 1069
QY 546 LeuAlaGluProArgMetSerAspAlaAlaMetValHisProSerGluValAlaAlaLys 565
DB 1070 CTTGCTGAACACACGATGTCAGATGCTGCCATGGTTTCAACCATCTGAGGTAGTGCCAAA 1129
QY 566 GlySerArgTrpLysAspLeuPheGluProGlyValArgArgAlaLeuLeuValGlyVal 585
DB 1130 GGTTCAAGTTGAAAGATTGTTTGAACCTGGAGTGGCGGTGCCCTGTTAGTCGGTGT 1189
QY 586 GlyIleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyValLeuTyrTyrThrProGln 605
DB 1190 GGAATTCAGATCCTTCAACAGTTTGTGGAATAAACGGTGTCTGTACTATACCCACAA 1249
QY 606 IleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSer 625
DB 1250 ATTCTTGAGCAAGCTGGTGTGGCAGTTATTCTTTCCAAATTTGGTCTCAGCTCGGCATCA 1309
QY 626 AlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAla 645
DB 1310 GCATCCATCTTGATCAGTTCTCTCACTACCTTACTAATGCTTCTTGCATTTGGCTTGCC 1369
QY 646 MetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuLeuGlyThrIleProIleLeu 665
DB 1370 ATGCTGCTTATGGATCTTTCCGGAAGAAGGTTTTTGTGCTGCTAGGCACAATCCCAATCTTG 1429
QY 666 IleAlaSerLeuValIleLeuValValSerAsnLeuIleAspLeuGlyThrIleAlaHis 685
DB 1430 ATAGCATCTCTAGTTATCTCTGTTGTGTCCAATCTTAATTGATTGGGTACACATAGCCCAT 1489
QY 686 AlaLeuLeuSerThrValSerValIleValTyrPheCysCysPheValMetGlyPheGly 705
DB 1490 GCTTTGCTCTCCACCATCAGTTTATCGTCTACTTCTGCTGCTTCTGTTATGGGATTGGT 1549
QY 706 ProIleProAsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIle 725
DB 1550 CCCATCCCCAACATTTTATGTGACAGAGATCTTTCCAAACCAGGGTTCGTGGCCCTCTGTATT 1609
QY 726 AlaIleCysAlaPheThrPheTrpIleGlyAspIleValThrTyrSerLeuProVal 745
DB 1610 GCCATTGTGCTTTTACATTCTGGATCGGAGATATCATGTCACCTACAGCTTCTCTGTG 1669
QY 746 MetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeuIle 765
DB 1670 ATGTCGAATGCTATTGGACTGCGCGGTGTTTTCAGCATATATGAGTCGTATGCTGTATT 1729
QY 766 SerPheValPheValPheLeuLysValProGluThrLysGlyMetProLeuGluValIle 785
DB 1730 TCCTTTGTGTTCTGTTCTTCTTAAAGTCCCTGAGACAAGGGGATGCCCTTCTGAGGTTATT 1789
QY 786 ThrGluPhePheAlaValGlyAlaLysGlnAlaAlaLysAla 800
DB 1790 ACCGAATCTTTTGCAGTTGGTGGCAAGCAAGCGGCTGCAAAAGCC 1834

RESULT 2

CC423708/c

LOCUS

CC423708

911 bp

DNA

linear

GSS 19-MAY-2003

DEFINITION PUHOR49TB ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTa499I02,
genomic survey sequence.
ACCESSION CC423708
VERSION CC423708.1 GI:30903798
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 911)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
TITLE Maize Genomics Consortium
JOURNAL Unpublished (2003)
COMMENT Other GSSs: PUHOR49TD
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.
FEATURES
source
1..911
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBTa499I02"
/clone_lib="ZM_0.6_1.0_KB"
/note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
Cot selected genomic DNA library"
ORIGIN
Alignment Scores:
Pred. No.: 1.71e-130 Length: 911
Score: 1399.00 Matches: 267
Percent Similarity: 98.90% Conservative: 2
Best Local Similarity: 98.16% Mismatches: 3
Query Match: 34.33% Indels: 0
DB: 8 Gaps: 0
US-10-051-909-32 (1-800) x CC423708 (1-911)
QY 270 SerGlyGluLeuSerLeuLeuGluGluValGlyLeuGluValGlyGlyAspThrSerIleGlu 289
:::|||||
Db 818 ACAGGTGATTGTCCTTCTTCGAGGGTTGGAGTTGGAGGAGACACTTCCATTGAA 759
QY 290 GluTyrIleGlyProAlaThrGluAlaAlaAspLeuValThrAspGlyAspLys 309
|||||
Db 758 GAGTACATCATTTGACCTGCCACCGAGGCGCGATGATCTTTGTACTGACGGTGATAAG 699
QY 310 GluGlnIleThrLeuTyrGlyProGluGluGlnSerTyrPileAlaArgProSerLys 329
|||||
Db 698 GAACAAATCACACTTTATGGCCCTGAAGAAGCCAGTCATGGATTGCTCGACCTTCCAG 639
QY 330 GlyProIleMetLeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsn 349
|||||
Db 638 GGACCCAGCATGCTTGAAGTGTGCTTCTTGTCATCTCGTCTATGGGAGCATGGTGAAC 579
QY 350 GlnSerValProLeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMet 369
|||||
Db 578 CAGAGTGTACCCCTTATGGATCCGATTGTGACACTTTTGTGATGTCTCCATGAGAAATATG 519
QY 370 ProGlnAlaGlyGlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSer 389
|||||
Db 518 CCTCAAGCTGGAGGAAGTATGAGGAGCACATGTTTCCAACTTTTGAAGTATGTTTCAGT 459
QY 390 ValThrAspGlnHisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAsp 409
|||||

Db 458 GTCACAGATCAGCATGCCAAAATGAGCAGTGGGTGAAGAGAAATCTTCATAGGATGAC 399
QY 410 GluGluTyrAlaSerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeu 429
|||||
Db 398 GAGGAGTACGCATCTGATGTTGCAAGAGGTGACTATGAGGACAAATCTCCATAGCCCATTG 339
QY 430 LeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGly 449
|||||
Db 338 CTGTCCAGGCAGGCAACAGGTGCGAAGGGAAGGACATTGTGCCACCATGGTCACCGTGA 279
QY 450 SerAlaLeuSerMetArgArgGlnThrLeuLeuGlyGluGlyAspGlyValSerSer 469
|||||
Db 278 AGTGTCTTGAGCATGAGAAGGCAAGCCCTCTTTAGGGGAGGTGGAGATGTTGTGAGCAGC 219
QY 470 ThrAspIleGlyGlyTyrGlnLeuAlaTyrLysTyrSerGluLysGluGlyGluAsn 489
|||||
Db 218 ACTGATATCGTGGGGATGGCAGCTTGCTTGGAAATGGTCAGAGAAGGAAGGTGAGAAT 159
QY 490 GlyArgLysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySer 509
|||||
Db 158 GGTAGAAAGGAAGTGGTTTCAAAGAGTCTACTTGCACCAAGAGGGAGTTCTCTGGCTCA 99
QY 510 ArgArgGlySerIleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPhe 529
|||||
Db 98 AGAAGGGGCTCAATTGTTTCACCTTCCCGTGGTGGCGATGTTCTTGAGGGTAGTGAGTTT 39
QY 530 ValHisAlaAlaLeuValSerGlnSerAlaLeu 541
|||||
Db 38 GTACATGCTGCTGCTTTAGTAAGTCAGTCAGCACTT 3
RESULT 3
BZ723844
LOCUS
DEFINITION BZ723844 845 bp DNA linear GSS 24-FEB-2003
PUCFI60TD ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTa13J23,
genomic survey sequence.
ACCESSION BZ723844
VERSION BZ723844.1 GI:28519500
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 845)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
TITLE Maize Genomics Consortium
JOURNAL Unpublished (2003)
COMMENT Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.
FEATURES
source
1..845
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBTa13J23"
/clone_lib="ZM_0.6_1.0_KB"
/note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
Cot selected genomic DNA library"
ORIGIN
Alignment Scores:
Pred. No.: 7.09e-129 Length: 845
Score: 1382.50 Matches: 274
Percent Similarity: 97.52% Conservative: 1

```
Best Local Similarity: 97.16%  Mismatches: 6
Query Match: 33.93%  Indels: 3
DB: 8  Gaps: 0
```

US-10-051-909-32 (1-800) x BZ723844 (1-845)

Qy	303	LeuValThrAspGlyAspLysGluGlnIleThrLeuTyrGlyProGluGluGlyGlnSer	322
Db	3	CTTGTTACTGACGGTGATAAGGAACAAATCACACTTTATGGGCTGAAGAAGGCCAGTCA	62
Qy	323	TrpIleAlaArgProSerLysGlyProIleMetLeuGlySerValLeuSerLeuAlaSer	342
Db	63	TGGATTGTCGACCTTCCAAGGGACCCAGCATGCTTGGAAAGTGTGCTTCTCTTGCACT	122
Qy	343	ArgHisGlySerMetValAsnGlnSerValProLeuMetAspProIleValThrLeuPhe	362
Db	123	CGTCATGGGAGCATGGTGAACACAGAGTGATCCCTTTATGGATCCGATTGTGACACTTTT	182
Qy	363	GlySerValHisGluAsnMetProGlnAlaGlyGlySerMetArgSerThrLeuPhePro	382
Db	183	GGTAGTGTCATGAGAATATGCCTCAAGCTGGAGGAAGTATGAGGAGCACATTGTTCCA	242
Qy	383	AsnPheGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTrpAspGlu	402
Db	243	AACTTTGGAAAGTATGTTTCAGTGTCAACAGATCAGCATGCTTAAAAAATGAGCAGTGGGATGAA	302
Qy	403	GluAsnLeuHisArgAspAspGluGluTyrAlaSerAspGlyAlaGlyGlyAspTyrGlu	422
Db	303	GAGAATCTTCATAGGGATGACGAGGAGTACGCATCTGATGGTGCAGGAGGTGACTATGAG	362
Qy	423	AspAsnLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIle	442
Db	363	GACAATCTCCATAGCCCATTCGTGTCCAGGCAGGCAACAGGTGCGGAAGGAAGGACATT	422
Qy	443	ValHisHisGlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeuGlyGlu	462
Db	423	GTGCACCATGGTCAACCCTGGAAGTGCTTTTGAGCATGAGAAGGCCAAGCCTCTTAGGGGAG	482
Qy	463	GlyGlyAspGlyValSerSerThrAspIleGlyGlyGlyTrpGlnLeuAlaTrpLysTrp	482
Db	483	GGTGGAGATGGTGTGAGCAGCACTGATATCGGTGGGGGATGGCAGCTTGCTTGGAAATGG	542
Qy	483	SerGluLysGluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyrLeuHis	502
Db	543	TCAGAGAAAGGAAGGTGAGAAATGGTAGAAAGGAAGGTGGTTTCAAAAGAGTCTACTTGCAC	602
Qy	503	GlnGluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGlyGlyAsp	522
Db	603	CAAGAGGGAGTTCCTGGCTCAAGAAAGGGGCTCAATTGTTTCACTTCCCGGTGTGGCGAT	662
Qy	523	ValPheGluGlySerGluPheValHisAlaAlaAlaLeuValSerGlnSerAlaLeuPhe	542
Db	663	GTTCCTGAGGGTAGTGAGTTGTACATGCTGCTGCTTTAGTAAGTCAGTCAGCACTTTTC	722
Qy	543	SerLysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisProSerGluVal	562
Db	723	TC-AAGGGTCTTGCTGAACA-CGCATGTCAGATGCTGCCATGTTTTCACCCATCTGAGGTA	780
Qy	563	AlaAlaLysGlySerArg-TrpLysAspLeuPheGluProGlyValArgArgAlaLeuLe	582
Db	781	GCTGCCAAAGGTTACGTTTGAAAGATTGTTTGAACCTGGAGTGGAGCGTGCCCTGTT	840
Qy	582	uVal 583	
Db	841	AGTC 844	

RESULT 4
BZ989776/c
LOCUS
DEFINITION
BZ989776 795 bp DNA linear GSS 25-MAR-2003
FUGDY07TD ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTA362B13,
genomic survey sequence.
ACCESSION
BZ989776
VERSION
BZ989776.1 GI:292229913

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source

GSS.
Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 795)
Whitelaw, C.A.; Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and
Bennetzen, J.
Maize Genomics Consortium
Unpublished (2003)
Other_GSSs: PUGDY07TB
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.
Location/Qualifiers
1..795
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBTa362B13"
/clone_lib="ZM_0.6 1.0 KB"
/note="Vector: PCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
cot selected genomic DNA library"

Db 318 GAGGAGTACGCATCTGATGGTGCAGGAGGTGACTATGAGGACAAATCTCCATAGCCCATG 259

Qy 430 LeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGly 449

Db 258 CTGTCAGGCAGGCAACAGGTGCGGAAGGAAGACATTGTCACCATGGTCCACCGTGA 199

Qy 450 SerAlaLeuSerMetArgArgGlnThrLeuLeuGlyGluGlyGlyAspGlyValSerSer 469

Db 198 AGTGCTTTGAGCATGAGAAGCGCAAGCCTCTTAGGGGAGGTGGAGATGGTGTGAGCAGC 139

Qy 470 ThrAspIleGlyGlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsn 489

Db 138 ACTGATATCGGTGGGGATGGCAGCTTGCTTGGAAATGGTCAGAGAAGGAGGTGAGAA 79

Qy 490 GlyArgLysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySer 509

Db 78 GGTAGAAAGGAAGGTGGTTTCAAAAGAGTCTACTTGCACCAAGAGGGAGTTCCCTAGCTCA 19

Qy 510 ArgArgGlySerIleVal 515

Db 18 AGAAGGGGCTCAATTGTT 1

RESULT 5

CNS0A3H4

LOCUS

DEFINITION

Arabidopsis thaliana Full-length cDNA Complete sequence from clone

GSLTSIL50ZE01 of Silique of strain col-0 of Arabidopsis thaliana

(thale cress).

ACCESSION

BX828912

VERSION

BX828912.1 GI:42461191

KEYWORDS

HTC; GSLT cDNA.

SOURCE

Arabidopsis thaliana (thale cress)

ORGANISM

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 1135)

Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C.,

Menard,M., Cruaud,C., Quetier,F., Searpelli,C., Schachter,V.,

Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.

Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:

A Combined Approach to Evaluate and Improve Arabidopsis Genome

Annotation

Unpublished

2 (bases 1 to 1135)

Genoscope.

Direct Submission

Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :

BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr

- Web : www.genoscope.cns.fr)

The sequences are based on single pass reads.

Life Technologies (a division of Invitrogen) members carried out

full-length libraries construction : Temple G.

Genoscope members carried out sequencing and annotation : Castelli

V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,

Schachter V., Weissenbach J., Salanoubat M.

URGV INRA : Clepet C., Caboche M.

Annotation is based on the June 2003 version of the Arabidopsis

genome released by MIPS (Munich Information center for Protein

Sequences). 5 prime and 3 prime are assembled with Phrap.

http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full

length

http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.

FEATURES

source

1..1135

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/strain="Col-0"

/db_xref="taxon:3702"

/clone="GSLTSIL50ZE01"

/tissue type="Silique"

/plasmid="pCMVSPORT_6"

complement(1..1135)

gene

ORIGIN

/gene="At4g35300"

Alignment Scores:

Pred. No.: 3,96e-117 Length: 1135

Score: 1268.50 Matches: 247

Percent Similarity: 82.82% Conservative: 47

Best Local Similarity: 69.58% Mismatches: 58

Query Match: 31.13% Indels: 3

DB: 3 Gaps:

US-10-051-909-32 (1-800) x CNS0A3H4 (1-1135)

Qy 445 HisGlyHisArgGlySerAlaLeuSerMetArgArg---GlnThrLeuLeuGlyGluGly 463

Db 1 CACACTGCTCATGGAACCTCTTTCTACCTTCAGACATGGAAGTCAAGTGCAGGAGCTCAA 60

Qy 464 GlyAspGlyValSerSerThrAspIleGlyGlyGlyTrpGlnLeuAlaTrpLysTrpSer 483

Db 61 GGGGAAGGAGCGGGTAGTATGGGGATTGGAGGTGGATGGCAAGTGGCATGGAATGGACG 120

Qy 484 GluLysGluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyrLeuHisGln 503

Db 121 GAAAGAGAAGATGAATCGGGACAGAAAGAGGTGGGTTTAAACGGATATACTTGCATCAA 180

Qy 504 GluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGlyAspVal 523

Db 181 GAAGGTTTCCCAGGATCTCGACGTGGCTCAATTGTTTTCATTGCTTGGTGTGATGGAACC 240

Qy 524 PheGluGlySerGluPheValHisAlaAlaAlaLeuValSerGlnSerAlaLeuPheSer 543

Db 241 GGTGAG---GCAGATTTGTACAAGCGTCTGCTTGGTTAGCCAAACAGCTCTTTATTCC 297

Qy 544 LysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisProSerGluValAla 563

Db 298 AAAGACCTTCTCAAAGAACATACAATTTGGTCTGCTATGGTATCCATCCGAA---ACA 354

Qy 564 AlaLysGlySerArgTrpLysAspLeuPheGluProGlyValArgArgAlaLeuVal 583

Db 355 ACTAAAGGGTCAATTTGGCATGATCTTTCATGATCCTGGAGTCAAGCGTGCATTAGTCGTA 414

Qy 584 GlyValGlyIleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyValLeuTyrThr 603

Db 415 GGAGTTGGACTTCAATACTTTCAGCAGTTCTCAGGCATCAACGGAGTTCTTACTACACA 474

Qy 604 ProGlnIleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSer 623

Db 475 CCGCAATCCTTGAGCAGCGGGTGTTCGGGATCCTACTATCGAACATGGGAGATTAGTTCT 534

Qy 624 AlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIleGly 643

Db 535 TCCTCAGCATCCTTACTATAAGTGCATTGACAAACCTTTGTGATGTTTACCTGCAATAGCT 594

Qy 644 PheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuGlyThrIlePro 663

Db 595 GTTCAATGAGGCTCATGATCTTTCTGGTTCGAAGGACCTTGCTTCTCACCACGATACCA 654

Qy 664 IleLeuIleAlaSerLeuValIleLeuValValSerAsnLeuIleAspLeuGlyThrLeu 683

Db 655 ATCCTGATAGCATCTCTATTGGTTTTAGTAATCTCAAATCTTGTTCACATGACAGCATT 714

Qy 684 AlaHisAlaLeuLeuSerThrValSerValIleValTyrPheCysCysPheValMetGly 703

Db 715 GTGCACGCGGTCTTATCAACCGTAAGCGTTGTGCTCTACTTCTGCTTCTTCGTGATGGGT 774

Qy 704 PheGlyProIleProAsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeu 723

Db 775 TTCGGTCTCTGCTCCAAACATCCTCTGTTTCAGAGATTTTCCAACTCGAGTCCGCGGAATC 834

Qy 724 CysIleAlaIleCysAlaPheThrPheTrpIleGlyAspIleIleValThrTyrSerLeu 743

Db 835 TGCATCGCCATCTGCGCACCTCACCTTCTGGATCTGTGACATAATCGTCACTTACAGTCTC 894

JOURNAL
COMMENT

Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bcccenter.fcav.unesp.br
Plate: 076 row: D column: 10
Seq primer: T7 Promoter Primer.
Location/Qualifiers
1..733
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCVPT2076D10"
/lab_host="DH10B"
/clone_lib="RT2"
/note="Organ: Root tips (0.3cm-long) from adult plants;
vector: pSport1; Site 1: Sali; Site 2: NotI; An
unidirectional cDNA library generated from [Root
tips (0.3cm-long) from adult plants]. cDNA was prepared
from polyA+ mRNA using Superscript Plasmid System Kit
(Invitrogen). The double-strand cDNAs were fractionated
in a sepharose CL-2B 40cm-columns and fragments sizing
between 0.8 and 1.5 Kb were directionally cloned into the
vector. Details of each source of RNA and library
construction can be obtained at
http://sucest.lad.ic.unicamp.br/public"

FEATURES
source

ORIGIN

Alignment Scores:
Pred. No.: 7 71e-106 Length: 733
Score: 1154.00 Matches: 221
Percent Similarity: 93.88% Conservative: 9
Best Local Similarity: 90.20% Mismatches: 14
Query Match: 28.32% Indels: 1
DB: 6 Gaps: 0

US-10-051-909-32 (1-800) x CA146057 (1-733)

QY 340 LeuAlaSerArgHisGlySerMetValAsnGlnSerValProLeuMetAspProIleVal 359
DB 1 CTCGCATCTCGTCATGCGCAGCATGGTGAAACCAGAGTGTAACCCCTTATGATCCGATTGTG 60
QY 360 ThrLeupheGlySerValHisGluAsnMetProGlnAlaGlyGlySerMetArgSerThr 379
DB 61 ACACATTTTGGGAGTGTCACAGAGATATGCTCCTCAAGCTGGAGGAAGTATGAGGAGCACA 120
QY 380 LeupheProAsnPhelGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGln 399
DB 121 TTGTTTCCAAACTTGGAAAGTATGTTTCAGTGTACACAGATCAGCATGCCAAAAATGAGCAG 180
QY 400 TrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAlaSerAspGlyAlaGlyGly 419
DB 181 TGGGACGAAGAGATCTTCACAGGGACGATGAGGAGTATGCATCTGATGTCAGGAGGT 240
QY 420 AspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGly 439
DB 241 GATTATGAGGACATCTCCACAGCCCATTTGCTCCAGGCAGACACAAAGTGTGAAGGA 300
QY 440 LysAspIleValHisHisGlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeu 459
DB 301 AAGGACATTTGTGCACCATGGTCACCGTGGAAAGTCTTTGAGCATGAGAGGCAAGCCTC 360
QY 460 LeuGlyGluGlyGlyAspGlyValSerSerThrAspIleGlyGlyGlyTrpGlnLeuAla 479
DB 361 TTGGGGGAGGCTGGAGAGGGTGTGAGCAGCACTGATATTGGTGGAGGATGGCAGCTTGCA 420
QY 480 TrpLysTrpSerGluLysGluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgVal 499

Db 421 TGGAAATGGTCAGAGAAGGAAGGTGAGGATGGTAAGAAGGAGGTCTCAAAGAGTC 480
QY 500 TyrLeuHisGlnGluGlyValProGlySerArgGlySerIleValSerLeuProGly 519
Db 481 TACTTACACCAAGAGGGAGTTCCTGGCTCAAGAATGGGCTCAATTGTTCACTTCCTGGT 540
QY 520 GlyGlyAspValPheGluGlySerGluPheValHisAlaAlaLeuValSerGlnSer 539
Db 541 GGTGGCAATGTTCTCTGAGGTGGCGAGTTGTACATGCTGCTGCTTGTAGTAAGCCCGTCA 600
QY 540 AlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisPro 559
Db 601 GCACATTTTCTCGAAGATCTTACCGAACCAACGATGTCTGTGTGCTGCTGCTTACCCCA 660
QY 560 SerGluValAlaLalysGlySerArg-TrpLysAspLeupheGluProGlyValArgAr 579
Db 661 TCCCAGGTAGTGCACAAAGGTTCAAAGTTGGAAGATTTGTTTGAACCTGGTGTGAGGCG 720
QY 579 gAlaLeuLeuVal 583
Db 721 TGCCCTTGTAGTC 733

RESULT 8
CA075335
LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 696)
Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
The libraries that made SUCEST
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bcccenter.fcav.unesp.br
Plate: 069 row: A column: 06
Seq primer: T7 Promoter Primer.

FEATURES
source

Location/Qualifiers
1..696
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCJFAM1069A06"
/lab_host="DH10B"
/clone_lib="AM1"
/note="Organ: Apical meristem and tissues surrounding of
mature plants; Vector: pSport1; Site_1: Sali; Site_2:
NotI; An unidirectional cDNA library generated from
[Apical meristem and tissues surrounding of mature
plants]. cDNA was prepared from polyA+ mRNA using
SuperScript Plasmid System Kit (Invitrogen). The
double-strand cDNAs were fractionated in a sepharose
CL-2B 40cm-columns and fragments sizing between 0.8 and
1.5 Kb were directionally cloned into the vector. Details
of each source of RNA and library construction can be
obtained at http://sucest.lad.ic.unicamp.br/public"

ORIGIN

Alignment Scores:
Pred. No.: 1.16e-104 Length: 696
Score: 1142.00 Matches: 214
Percent Similarity: 95.22% Conservativeness: 5
Best Local Similarity: 93.04% Mismatches: 11
Query Match: 28.02% Indels: 0
DB: 6 Gaps: 0

US-10-051-909-32 (1-800) x CA075335 (1-696)

QY 305 ThrAspGlyAspLysGluGlnIleThrLeuTyrGlyProGluGluGlyGlnSerTrpIle 324
Db 1 ACTGATGTTGATAAGGAACAAATCACACTTATGGCCCTGAAGAAGGCCAGTCATGGATT 60
QY 325 AlaArgProSerLysGlyProIleMetLeuGlySerValLeuSerLeuAlaSerArgHis 344
Db 61 GCTCGACCTTCCAAAGGGACCCAGCATGCTTGAAGTGTGCTGCTCTCGCATCTCGTCAT 120
QY 345 GlySerMetValAsnGlnSerValProLeuMetAspProIleValThrLeuPheGlySer 364
Db 121 GGCAGCATGTTGAACACAGAGTGTACCCCTTATGGATCCGATTGTGACACTTTTTCGGAGT 180
QY 365 ValHisGluAsnMetProGlnAlaGlyGlySerMetArgSerThrLeuPheProAsnPhe 384
Db 181 GTCCACGAGAATATGCTCAAGCTGGAGGAAGTATGAGGAGCACATTTTCCAACTTT 240
QY 385 GlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTrpAspGluAsn 404
Db 241 GGAAGTATGTTCAAGTGTACAGATCAGCATGCCCCAAATGAGCAGTGGGACGAGAGAAT 300
QY 405 LeuHisArgAspAspGluGluTyrAlaSerAspGlyAlaGlyGlyAspTyrGluAspAsn 424
Db 301 CTTCAAGGACGATGAGGAGTATGATCTGATGTTGAGGAGGTGATTATGAGGACAAT 360
QY 425 LeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIleValHis 444
Db 361 CTCCACAGCCCATTTGCTTCCAGGACAGACAACAGTATGGAAGGAAGGACATTTGTGCAC 420
QY 445 HisGlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeuGlyGluGlyGly 464
Db 421 CATGTCACCGTGAAGTCTTTGAGCATGAGAGGCAAGGCCTCTTGGGGGAGGCTGGA 480
QY 465 AspGlyValSerSerThrAspIleGlyGlyGlyTrpGlnLeuAlaTrpLysTrpSerGlu 484
Db 481 GAGGTTGTGAGCAGCATGATATTGTTGGAGGATGGCAGCTTGCATGGAATGGTCAGAG 540
QY 485 LysGluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyrLeuHisGlnGlu 504
Db 541 AAGGAAGGTGAGGATGTTAAGGAAGGTGTTTCAAAAGAGTCTACTTACACCAAGAG 600
QY 505 GlyValProGlySerArgArgGlySerIleValSerLeuProGlyGlyGlyAspValPhe 524
Db 601 GGAGTTCCTGNCCTCAAGAAATGGGCTCANTNGTNTCACTTCTGTTGGTGGCGATGTTCTCT 660
QY 525 GluGlySerGluPheValHisAlaAla 534
Db 661 GAGGTTGGCGAGTTNGTACATGCTGCTGCT 690

RESULT 9
CK208781
LOCUS 1082 bp mRNA linear EST 08-DEC-2003
DEFINITION FGAS020500 Triticum aestivum FGAS: Library 5 GATE 7 Triticum
aestivum cDNA, mRNA sequence.
CK208781
ACCESSION CK208781.1 GI:39571171
VERSION EST.
KEYWORDS Triticum aestivum (bread wheat)
SOURCE Triticum aestivum
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.
REFERENCE 1 (bases 1 to 1082)

AUTHORS
TITLE
JOURNAL
COMMENT

Allard, F., Crosby, W.L., Danyluk, J., Eudes, F., Frick, M., Gaudet, D.,
Genswein, B., Graf, R., Gulick, P., Hrycan, L.D., Laroche, A.,
Links, M.G., McCarthy, E.L., Monroy, A., Muzak, I., Nilsson, D.,
Penniket, C., Roach, J.L. and Sarhan, F.
Functional Genomics of Abiotic Stress In Wheat and Canola Crops
Unpublished (2003)
Contact: Wm L Crosby
Bioinformatics
University of Saskatchewan, Department of Computer Science
1C101 Engineering Building, 57 Campus Drive, Saskatoon,
Saskatchewan, S7N 5A9, Canada
Tel: 306 966 1769
Fax: 306 966 2033
Email: fgas_est@cs.usask.ca

This sequence is the direct result of the Base calling software
Phred (default parameters). It is the raw base calls. To aid in the
identification of the high quality insert the software Lucy
(default parameters) has been run on this sequence. Lucy identified
the region [31,846].

Plate: L5B015 row: E column: 19.

FEATURES
source

Location/Qualifiers
1..1082
/organism="Triticum aestivum"
/mol_type="mRNA"
/db_xref="taxon:4565"

/clone_lib="Triticum aestivum FGAS: Library 5 GATE 7"
/note="Vector: pCMV.SPORT6; Crown and developmental stages
of spike formation in wheat cultivar Norstar. 4 mRNA
populations were combined before constructing the library.
The first mRNA population is from 1cm crown sections after
30 days of cold acclimation. The second is from 1cm crown
sections after 11 days of deacclimation (before
deacclimation plants were fully vernalized for 49 days).
The third is from different developmental stages of spike
formation (5 to 50mm) that still have not emerged from the
leaf (dissection required). The last is from different
developmental stages of spike and seed formation after
having emerged from the leaf (visible). First strand
synthesis in this library was done in the presence of
methylated dCTP thereby protecting from internal cleavage
with NotI."

ORIGIN

Alignment Scores:
Pred. No.: 3.98e-102 Length: 1082
Score: 1120.00 Matches: 242
Percent Similarity: 81.23% Conservativeness: 35
Best Local Similarity: 70.97% Mismatches: 49
Query Match: 27.48% Indels: 20
DB: 7 Gaps: 4

US-10-051-909-32 (1-800) x CK208781 (1-1082)

QY 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrp 75
Db 65 ATGTGGGGCGCGTGTGTCGTCATAGCGCGCTCCATCGGAACCTGCTGAGGGGTGG 124
QY 76 AspAsnAlaThrIleAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsn 95
Db 125 GATAATCGACCATTCAGGTGCGCTCCTGTACATAAAGAGAGTTTCAGCTGGAGACC 184
QY 96 GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115
Db 185 CAGCCCTTGATCGAGGGCTCATCGTGGCCATGCTCATCGGGGGCGGCGTTATCACC 244
QY 116 ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer 135
Db 245 AGTTCTCCGGGGCGGTGCTGATGCTGTTGGTAGGGCGCCCTGCTGATCGCTCTCT 304
QY 136 IleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyrValLeuLeu 155
Db 305 GTCCTCTACTTTGTCTAGTGGCCTGGTGTGATGCTCTTGGCGCCCAACGCTCTATGTTGCTC 364

```
QY 156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr 175
Db 365 TTGGCCAGGTCATCGACGGTTCGGTATCGGTTTGGCTGTGCACCCCTGTCCCTCTTAC 424

QY 176 IleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSer 195
Db 425 ATTTCGAGACCGCCCGACTGACATTAGAGGGCTGCTGAACACGCTGCCGAGTTCACT 484

QY 196 GlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSer 215
Db 485 GGGTCAGGAGGGATGTTCTTTCTTACTGCATGGTGTTCACCATGTCGCTCATGCCGAG 544

QY 216 ProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeu 235
Db 545 CCTGACTGGAGAATCATGCTTTGGGGTTTGTTCGATCCCGTCGCTTATGTAATTTGCCATTG 604

QY 236 ThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGlu 255
Db 605 ACTGTCTTCTATTGCCCCGAGTCGCCGAGATGGCTTGTGAGCAAGGAAGAAATGGCCGAG 664

QY 256 AlaLysLysValLeuGlnLysLeuArgGlyLysAspAspValSerGlyGluLeuSerLeu 275
Db 665 GCCAAGCGAGTGTGCAAGGACTCGCGGGAAGGAAGATGTCTCAGGAGAAATGGCCCTT 724

QY 276 LeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIleGlyPro 295
Db 725 CTTGTTGAAGGATGGGTGTTGGAAAGACACACATTTTGAGGAATACATAATTGGGCCCT 784

QY 296 AlaThrGluAlaAlaAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyr 315
Db 785 GATGATGAGCTTGCTGATGACGGTCTGGTCCA---GATGAAGAGAA-GTGAAGCTGTAC 840

QY 316 GlyProGluGlyGlnSerTyrIleAlaArgProSerLysGlyProIleMetLeuGly 335
Db 841 GGAGCTGAAGAGGGGTA-TCTTGGATCGCCCGT-----CTGTTAGGG 881

QY 336 SerVal-----LeuSerLeuAlaSerArgHisGlySerMetVal 348
Db 882 GCGGTGGCAAAGTGCACCTTGAAGCGCTTGGGTCTCATGCTCGTCATGGGAGTATGGTT 941

QY 349 AsnGlnSerValProLeuMetAspProIleValThrLeuPheGlySerValHisGluAsn 368
Db 942 AGTCAGGGTAA-TCTCTCGTGGACCCACT-GTTACTCTCTTCGGAAGTGCCATAGAAGA 999

QY 369 MetProGlnAlaGlyGlySerMetArgSerThrLeuPheProAsnPheGlySerMetPhe 388
Db 1000 ---TCCCTGAGTAAATGGGGAGCATTCGCCCAATGTTTCTTAACCTT-GGCCACATGTTT 1055

QY 389 Ser 389
Db 1056 ACG 1058

RESULT 10
BQ752737
LOCUS
DEFINITION
WHE4118_E11_I22ZS Wheat salt-stressed root cDNA library Triticum aestivum cDNA clone WHE4118_E11_I22, mRNA sequence.

ACCESSION
BQ752737
VERSION
BQ752737.1
KEYWORDS
GI:21930519
SOURCE
Triticum aestivum (bread wheat)
ORGANISM
Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae; Triticeae; Triticum.
REFERENCE
1 (bases 1 to 829)
Anderson,O.D., Akhunov,E., Chao,S., Crossman,C., Deal,K., Dvorak,J., Lazo,G.R., Pham,J., Rausch,C.J., Wilson,C. and Woo,J.
The structure and function of the expressed portion of the wheat genomes - Salt-stressed root cDNA library
Unpublished (2002)
JOURNAL
Contract: Olin Anderson
COMMENT
US Department of Agriculture, Agriculture Research Service, Pacific
```

```
West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818
Email: candersn@pw.usda.gov
Sequences have been trimmed to remove vector sequence and low
quality sequence with phred score less than 20
Seq primer: SK primer.

FEATURES
            Location/Qualifiers
     1..829
        /organism="Triticum aestivum"
        /mol_type="mRNA"
        /cultiivar="Chinese Spring"
        /db_xref="taxon:4565"
        /clone="WHE4118_E11_I22"
        /tissue_type="Roots"
        /dev_stage="Full tillering"
        /lab_host="E. coli SOLR"
        /clone_lib="Wheat salt-stressed root cDNA library"
        /note="Vector: Lambda Uni-ZAP XR, excised phagemid
pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; Hydroponic
plants grown to full tillering stage were treated with 150
mM NaCl for either 12 hours or 7 days. Root tissues of the
plants subjected to both types of treatment were collected
separately at University of California, Davis (E. Akhunov
and K. Deal in J. Dvorak's lab). Total RNA was prepared
separately from the two samples (12h and 7day treatments),
and equal amount of RNA was then pooled. PolyA RNA was
purified from the pooled RNA, a cDNA library was made, and
the cDNA clones were in vivo excised to give pBluescript
SK(-) phagemids in J. Dvorak's lab (E. Akhunov, J. Dvorak)
at the University of California, Davis. Colony plating,
plasmid DNA preparations and DNA sequencing were performed
in the OD Anderson lab (all other authors)."
```

ORIGIN

```
Alignment Scores:
Pred. No.:      2.18e-100      Length:      829
Score:          1101.00       Matches:    220
Percent Similarity: 89.89%    Conservative: 20
Best Local Similarity: 82.40%  Mismatches:  25
Query Match:    27.02%       Indels:      2
DB:              5           Gaps:        2
```

US-10-051-909-32 (1-800) x BQ752737 (1-829)

```
QY 486 GluGlyGluAsnGlyArgLysGlyGlyPheLysArgValTyrLeuHisGlnGluGly 505
Db 28 AAAGCGGAGGATGGCAAGAAGGAGGAGGCTTCAAAAAGATCTACTTGCCACCAAGCGG 87

QY 506 ValProGlySerArgArgGlySerIleValSerLeuProGlyGlyAspValPheGlu 525
Db 88 GTGCCCGACTCAAGAAGGGGCTCTGTTGTTTTCATCTCCTGGTGGGGTGATGCCACCAA 147

QY 526 ---GlySerGluPheValHisAlaAlaLeuValSerGlnSerAlaLeuPheSerLys 544
Db 148 GGGGSCAGTGGGTTTATACACGCTGCTGCTTTGGTAAGCCACTCGGCTCTTTACTCCAAG 207

QY 545 GlyLeuAlaGluProArgMet---SerAspAlaAlaMetValHisProSerGluValAla 563
Db 208 GATCTTATGAAGAGCGGTATGGCGCGCGGTCCAGCCCATGATCCATCCATTGGAGGCAGCT 267

QY 564 AlaLysGlySerArgTrpLysAspLeuPheGluProGlyValArgArgAlaLeuVal 583
Db 268 CCCAAAGGTTCAATCTGGAAGATCTGTTTGAACCTGGTGTGAGCGGTGATGTTTCGTC 327

QY 584 GlyValGlyIleGlnIleLeuGlnPheAlaGlyIleAsnGlyValLeuTyrThr 603
Db 328 GGCCTTGAATTCAGATGCTTCAGCAGTTTGTGGAATAAATGGAGTTCTCTACTATACT 387

QY 604 ProGlnIleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSer 623
Db 388 CCTCAAAATCTGGAGCAAGCTGGTGGCGGGTCTCTTCTTCCAAATCTTGGCCTCAGTTCA 447
```

```
QY 624 AlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIleGly 643
Db 448 GCATCAGCGTCCATCTTGATCAGCTCTCTCACCACCTTACTCATGCTCCCAAGCATTTGT 507
QY 644 PheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuLeuGlyThrIlePro 663
Db 508 GTAGCCATGAGACTTATGGATATATCTGGAAGAAGTTTCTTCTACTGGGCACAATTCCC 567
QY 664 IleLeuIleAlaSerLeuValIleLeuValValSerAsnLeuIleAspLeuGlyThrLeu 683
Db 568 ATCTTGATAGCATCCCTAATTGTTTGGTGTGTCTCAATGTTATCAACTTGAGTACGGTG 627
QY 684 AlaHisAlaLeuLeuSerThrValSerValIleValTyrPheCysCysPheValMetGly 703
Db 628 CCCACGCTGTCTCTCCACAGTTAGGCTCATTTGTCTACTTGTCTTGTCTATGGGC 687
QY 704 PheGlyProIleProAsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeu 723
Db 688 TTTGGCCCAATCCCCAACATTTCTATGTGCAGAGATTTTCCCCACCAGAGTCCGTGGTGT 747
QY 724 CysIleAlaIleCysAlaPheThrPheThrIleGlyAspIleIleValThrTyrSerLeu 743
Db 748 TGCATCGCTATTGGCCCTCACATCTGGATATGTGACATTATTGTTACCTACAGCCTG 807
QY 744 ProValMetLeuAsnAlaIle 750
Db 808 CCTGTGATGCTGAATGCTATT 828
```

```
RESULT 11
CA170109
LOCUS CA170109 871 bp mRNA linear EST 24-SEP-2003
DEFINITION SCQGSB1083B09.g SB1 Saccharum officinarum cDNA clone SCQGSB1083B09
5', mRNA sequence.
ACCESSION CA170109
VERSION CA170109.1 GI:35093205
KEYWORDS EST.
SOURCE Saccharum officinarum
ORGANISM Saccharum officinarum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
complex.
1 (bases 1 to 871)
Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
The libraries that made SUCEST
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bcccenter.fcav.unesp.br
Plate: 083 row: B column: 09
Seq primer: T7 Promoter Primer.
```

```
FEATURES
source
1..871
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCQGSB1083B09"
/lab_host="DH10B"
/clone_lib="SB1"
/note="Organ: Stalk Bark from adult plants; Vector:
psport1; Site 1: SalI; Site 2: NotI; An unidirectional
cDNA library generated from [Stalk Bark from adult
plants]. cDNA was prepared from polyA+ mRNA using
SuperScript Plasmid System Kit (Invitrogen). The
double-strand cDNAs were fractionated in a sepharose
```

CL-2B 40cm-columns and fragments sizing between 0.8 and 1.5 Kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at <http://sucst.lad.ic.unicamp.br/public>

ORIGIN

```
Alignment Scores:
Pred. No.: 2.98e-100 Length: 871
Score: 1100.00 Matches: 215
Percent Similarity: 86.11% Conservative: 33
Best Local Similarity: 74.65% Mismatches: 40
Query Match: 26.99% Indels: 2
DB: 6 Gaps: 0
US-10-051-909-32 (1-800) x CA170109 (1-871)
QY 420 AspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGly 439
Db 9 GACTATGAAACAATGTACACAGCCCTCTGCTGTCCCGACAGAATACAAAGTCCCGAAGCG 68
QY 440 LysAspIleValHisHisGlyHisArgGlySerAlaLeuSerMetArgGlnThrLeu 459
Db 69 AAGGGCATTCACATCACGGGCGCCGTGGAAGTCTTTGGGTTTGAGAAGAAGAGCTG 128
QY 460 LeuGlyGluGlyGlyAspGlyValSerSerThrAspIleGlyGlyTyrGlnLeuAla 479
Db 129 TCCGATGAGGGTGGTGAGGCGAGCAACCAGCACCTGGCATTGGTGGCGATGGCAGCTGCC 188
QY 480 TrpLysTrpSerGluLysGluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgVal 499
Db 189 TGGAATGGTCAGAGCGGAGAGGTGAGGACGGTAAGAGGAAGGGAGTTTCAAGAAGATC 248
QY 500 TyrLeuHisGlnGluGlyValProGlySerArgArgGlySerIleValSerLeuProGly 519
Db 249 TACTTACCAAGAGGGAGTTGCTGGCTCAAGAGGGGATCTGTGTCTCACTTCTCTGGT 308
QY 520 GlyGlyAspValPheGluGlySerGluPheValHisAlaAlaLeuValSerGlnSer 539
Db 309 GGAGGTGATGCTCTGAAGGTGGTAAGTTTATACATCTGCTGCTCTGGTCAAGCAGTCA 368
QY 540 AlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisPro 559
Db 369 GCTCTTTACCCGAGGGATATTACAGAACAGCGCATGGCCGGTCCAGCTACAATGCACCCA 428
QY 560 SerGluValAlaAlaLysGlySerArgTrpLysAspLeuPheGluProGlyValArgArg 579
Db 429 TCAGAGGCAGCTGCCAAAGTTCCAAAGCTGGAGAGACCTGTTTGAACCTGGTGTGAGACGT 488
QY 580 AlaLeuLeuValGlyValGlyIleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyVal 599
Db 489 GCCCTGTTGGTTGTTGTAATTCAGATCCCTTCAACAGATTTCGCGGGAATAAATGGNGTT 548
QY 600 LeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPhe 619
Db 549 CTCTACTACACTCCGCAATATTATGGAGCAAGCTGGTGGCAGTTTCTTATTTCCAATCTT 608
QY 620 GlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeu 639
Db 609 GGTCCTCAGTTCCGATCAGCATCCATCCCTAATTAGTCTGTCACTGCCCTACTTATGCTN 668
QY 640 ProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuLeu 659
Db 669 CCTAGCATTTGGTTTAGCCATGAGACTTATGGACGTATCTTGGAGAAGGTTTCTGTGCTA 728
QY 660 GlyThrIleProIleLeuIleAlaSerLeuValIleLeuValValSerAsnLeuIleAsp 679
Db 729 AGCACAATTCAGTCTGTAGTATCTTTAATTGCTCTGCTGCTGCTAATGTTATCGAG 788
QY 680 LeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyrPheCysCys 699
Db 789 TTGGGTACAGTGGTCCATGCTGTGCTCTCCACATT-AGTGTCAATCACCTACTTATGCTGC 847
QY 700 PheValMetGlyPheGlyProIle 707
```


Db 848 TTCAA-AATGGCTTTGGCCCCATT 870
RESULT 12
BX841312 1107 bp mRNA linear EST 11-FEB-2004
LOCUS BX841312 Arabidopsis thaliana Flowers and buds Col-0 Arabidopsis
DEFINITION thaliana cDNA clone GSLTFB382G07 SPRIM, mRNA sequence.
ACCESSION BX841312
VERSION BX841312.1 GI:42535395
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi
REFERENCE 1 (bases 1 to 1107)
AUTHORS Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C.,
Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,
Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
TITLE Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
JOURNAL Unpublished (2004)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequenage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS.(Munich Information center for Protein
Sequences).
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/EST
http://www.genoscope.cns.fr/cgi-bin/ggb?source=Arabidopsis.
FEATURES
source
1..1107
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecotype="Col-0"
/db_xref="taxon:3702"
/clone="GSLTFB382G07"
/tissue_type="Flowers and buds"
/clone_lib="Arabidopsis thaliana Flowers and buds Col-0"
ORIGIN
Alignment Scores:
Pred. No.: 1.14e-98 Length: 1107
Score: 1086.00 Matches: 220
Percent Similarity: 77.27% Conservative: 52
Best Local Similarity: 62.50% Mismatches: 71
Query Match: 26.65% Indels: 9
DB: 5 Gaps: 4
US-10-051-909-32 (1-800) x BX841312 (1-1107)
Qy 41 SerSerGlnGluProValThrSerAspIleLeuGluAspLysMetSerGlyAlaVal 60
Db 24 TCCTCGTAATTGCCGGCGGAATTCTACTAAAGCAGTTTGTGTGAGCAATGAAGGAGCGACT 83
Qy 61 LeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrpAspAsnAlaThrIle 80
Db 84 CTCGTGCTCTCGCCGCCACAATCGGCAATTTCTTACAAGGATGGACAATGCCACCATT 143
Qy 81 AlaAlaAlaValLeuTyrlleLysLysGluPheGlnLeuGlnAsnGluProThr---Val 99
Db 144 GCTGGAGCTATGTTTATATCAACAAAGACTTGAATCTA-----CCAACCTCTGTT 194

Qy 100 GluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThrPhSerGly 119
Db 195 CAAGGTCTTGTGCTGCTATGTGATTGATCGGTGCAACGGTCATCAGACTTGCTCAGGA 254
Qy 120 ProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSerIleLeuTyrlPhe 139
Db 255 CCGATACTGTGATTGGCTCGGCAGACGCCCATGTCTCATTTTATCATCAGTTATGTATTC 314
Qy 140 PheSerGlyLeuIleMetLeuTrpSerProAsnValTyrlleLeuLeuAlaArgPhe 159
Db 315 GTCTCGGTTTGATAATGTTGGTCTCCCAATGCTATGTCTGTGCTTTGCTAGGCTT 374
Qy 160 ValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyrlleSerGluIle 179
Db 375 CTTAATGGGTTTGGTGC CGGGCTCGCGGTTACACTTGTCCCTGTGTTACATTTCTGAAACC 434
Qy 180 AlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSerGlySerGlyGly 199
Db 435 GCTCCTCCGAGATCAGAGGACAGTTAAATACTCTCCCTCAGTTTCTTTGGCTCTGGTGA 494
Qy 200 MetPheLeuSerTyrlCysMetValPheGlyMetSerLeuSerProSerProAspTrpArg 219
Db 495 ATGTTTTGTCTACTACTGTATGTTTTCACATATGTCCCTGAGTACCTCCCTAGCTGGAGA 554
Qy 220 IleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeuThrIlePheTyrl 239
Db 555 GCCATGCTCGGTGCTCTCGATCCCTTCTCTTTATTATTGTTTCTCACGGTGTTTTAT 614
Qy 240 LeuProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGluAlaLysLysVal 259
Db 615 TTGCCCGAGTCTCCTCGTTGGCTGTTAGTAAAGGAAGATGGACGAGGCTAAGCGAGTT 674
Qy 260 LeuGlnLysLeuArgGlyLysAspValSerGlyGluLeuSerLeuLeuGluGly 279
Db 675 CTTCAACAGTTATGTGGCAGAGAAGATGTTACCGATGAGATGGCTTTACTAGTTGAAGGA 734
Qy 280 LeuGluValGlyGlyAspThrSerIleGluGluTyrlleIleGlyProAlaThrGluAla 299
Db 735 CTAGATATAGGAGGAGAAAAACAATGGAAGATCTCTTAGTAACCTTTGGAGGATCATGAA 794
Qy 300 AlaAspAspLeuValThrAspGlyAspLysGlu--GlnIleThrLeuTyrlGlyProGlu 318
Db 795 GGTGATGATACACTTGAAACCGTTGATGAGGATGGACAATGCGGCTTTATGGAACCCAC 854
Qy 319 GluGlyGlnSerTrpIleAlaArgPro-SerLysGlyProIleMetLeuGlySerValle 338
Db 855 GAGAATCAATCGTACCTTGTCTAGACCTGTCCCAGAACAAAATA-----GCTCACTT 905
Qy 338 uSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValProLeuMetMetAspProI 358
Db 906 GGGGCTACGCTCTCGCCACGGAAGCTTAGCAAAACCAAGCATGATCCTTAAAGATCCGCT 965
Qy 358 eValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGlyGlySerMetArgSe 378
Db 966 CGTCAATCTTTTGGCAGTCTCCACGAGAAGATGCCAGAGCGGCGGAAAACTCGGAG 1025
Qy 378 rThrLeuPheProAsnPheGlySerMetPheSer 389
Db 1026 TGGGATTTTCCAACATTTTCGGGAGCATGTTTCAGT 1059
RESULT 13
CF445664
LOCUS EST682009 normalized cDNA library of onion Allium cepa cDNA clone
DEFINITION ACAIL96, mRNA sequence.
ACCESSION CF445664
VERSION CF445664.1 GI:34468366
KEYWORDS EST.
SOURCE Allium cepa (onion)
ORGANISM Allium cepa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Alliaceae;
Allium.

(Invitrogen). The double-strand cDNAs were fractionated in a sepharose CL-2B 40cm-columns and fragments sizing between 0.8 and 1.5 Kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at <http://sucest.lad.ic.unicamp.br/public>

ORIGIN

Alignment Scores:
Pred. No.: 1.21e-96 Length: 769
Score: 1063.50 Matches: 224
Percent Similarity: 92.09% Conservative: 9
Best Local Similarity: 88.54% Mismatches: 17
Query Match: 26.10% Indels: 3
DB: 6 Gaps: 1

US-10-051-909-32 (1-800) x CA243478 (1-769)

Qy 530 ValHisAlaAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluPro 549
Db 2 GTACATGCTTCTGCTTTAGTAAGCCAGTCAGCACTTCTCGAAGGATCTTACCGAACCA 61
Qy 550 ArgMetSerAspAlaAlaMetValHisProSerGluValAlaAlaAlaLysGlySerArgTrp 569
Db 62 CGCATGTCTGGTCTGCCATGGTTTCACCCATCTGAGGTAGTGCCTAAAGGTTCAAGTTGG 121
Qy 570 LysAspLeuPheGluProGlyValArgArgAlaLeuValGlyValGlyIleGlnIle 589
Db 122 AAAGATTGTTTGAACCTGGTGTGAGCGGTGCCCTGTGTAGTCGGTGTGGAATTCAGATC 181
Qy 590 LeuGlnGlnPheAlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGln 609
Db 182 CTTCAACAGTTTGTCTGGAATAAATGGTGTCTGTACTATACCCCAAAATTCGAGCAA 241
Qy 610 AlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeu 629
Db 242 GCTGGCGTGGCAGTTCTCTTCCAAATCTTGGTCTCAGCTCGGCATCAGCATCCATCTTG 301
Qy 630 IleSerSerLeuThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMet 649
Db 302 ATCAGTTCTCTCACTACCTTACTGATGCTTCCTAGCATTTGGCTTAGCCATGAGACTCATG 361
Qy 650 AspLeuSerGlyArgArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLeu 669
Db 362 GATCTTTCTGGAAGAAGGTTTGTCTGCTAGGCACAATTCCAATCTTGATGATCATCTTTA 421
Qy 670 ValIleLeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSer 689
Db 422 GTTATCCTGGTCTGTCTCCAATGTTATTGACTTGGGTACAGTGGCCCATGCTGGCTCTCC 481
Qy 690 ThrValSerValIleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsn 709
Db 482 ACAGTCAGTGTATCATCTACTTCTGCTGTTTGTGATGGGATTTGGTCCCATCCCCAAC 541
Qy 710 IleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAla 729
Db 542 ATTCTATGTGCAGAGATCTTTCCAACTAGGTTTCGCGGTCTCTGCATTTGCCATCTGTGCC 601
Qy 730 PheThrPheTrp-IleGlyAspIleIleValThrTyrSer-LeuProValMetLeuAsnA 749
Db 602 TTGACATTTTGTGATCGGAGACATCATTTGCACCTACAGCCCTTCTCTGTGATGCTGAATG 661
Qy 749 laIleGlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeuIleSerPhe---V 768
Db 662 CTATTGGACTAACAGGTGTTTGGCATATATGCAATGATGCTTGAATGCGCTTTGGGG 721
Qy 768 alpHeValPheLeuLysValProGluThrLysGly 779
Db 722 TTTGGCTTCCCTTAAGGGTCTCTTGAGACAAAGGG 756

RESULT 15
CA067439 994 bp mRNA linear EST 23-SEP-2003
LOCUS

DEFINITION SCQGAD1065H09.g AD1 Saccharum officinarum cDNA clone SCQGAD1065H09
5', mRNA sequence.
ACCESSION CA067439
VERSION CA067439.1 GI:34918963
KEYWORDS EST.
SOURCE Saccharum officinarum
ORGANISM Saccharum officinarum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
complex.
REFERENCE 1 (bases 1 to 994)
AUTHORS Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
TITLE The libraries that made SUCEST
JOURNAL Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
COMMENT Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
<http://www.bcccenter.fcav.unesp.br>
Plate: 065 row: H column: 09
Seq primer: T7 Promoter Primer.
Location/Qualifiers
1 :994
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCQGAD1065H09"
/lab_host="DH10B"
/clone_lib="AD1"
/note="Organ: seedlings inoculated with Gluconacetobacter
diazotrophicans; Vector: pSport1; Site_1: SalI; Site_2:
NotI; An unidirectional cDNA library generated from
[seedlings inoculated with Gluconacetobacter
diazotrophicans]. cDNA was prepared from polyA+ mRNA using
SuperScript Plasmid System Kit (Invitrogen). The
double-strand cDNAs were fractionated in a sepharose
CL-2B 40cm-columns and fragments sizing between 0.8 and
1.5 Kb were directionally cloned into the vector. Details
of each source of RNA and library construction can be
obtained at <http://sucest.lad.ic.unicamp.br/public>

ORIGIN

Alignment Scores:
Pred. No.: 1.33e-95 Length: 994
Score: 1055.00 Matches: 205
Percent Similarity: 90.13% Conservative: 5
Best Local Similarity: 87.98% Mismatches: 21
Query Match: 25.89% Indels: 2
DB: 6 Gaps: 1

US-10-051-909-32 (1-800) x CA067439 (1-994)

Qy 272 GluLeuSerLeuLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyr 291
Db 1 GAATTGTCCTTCTTGTGTAAGGTTGGAGGTTGGAGGAGACACTTCGATTGAAGAGTAC 60
Qy 292 IleIleGlyProAlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGln 311
Db 61 ATCATTTGGCCCTGCCACTGAGGAGCGCATGATCATGTACTGATGGTATAGGACAA 120
Qy 312 IleThrLeuTyrGlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyPro 331
Db 121 ATCACACTTATGGCCCTGAAGAAGGCCAGTTCATGGATTGCTCGACCTTCCAAAGGACCC 180
Qy 332 IleMetLeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSer 351
Db 181 AGCATGCTTGAAGTGTGCTTTCTCTCGCATCTCTCGTATGGCAGCATGGTGAACACAGT 240

QY 352 ValProLeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGln 371
Db 241 GTACCCCTTTATGGATCCGATTGTGACACTTTTGGGAGTGTCCACGAGATATGCTCAA 300
QY 372 AlaGlyGlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThr 391
Db 301 GCTGGAGGAAGTATGAGGAGCACATTGTTCCAAACTTTGGAAGTATGTTTCAGTGTACA 360
QY 392 AspGlnHisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGlu 411
Db 361 GATCAGCATGCCAAAATGAGCAGTGGGACGAAGAGAAATCTTCACAGGGACGATGAGGAG 420
QY 412 TyrAlaSerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSer 431
Db 421 TATGCATCTGATGGTGCAGGAGGTGATTATGAGGACAAATCTCCACAGCCCATTTGCTGTCC 480
QY 432 ArgGlnAlaThrGlyAlaGluGlyLysAspIleValHisGlyHisArgGlySerAla 451
Db 481 AGGCAGACAAACAAGTGTGGAAGGAAAAGACATTGTGCACCATGTTCAACCGTGGAAAGTCT 540
QY 452 LeuSerMetArgArgGlnThrLeuLeuGlyGluGlyGlyAspGlyValSerSerThrAsp 471
Db 541 TTGAGCATGAGAAGGCAAAGCCTCTTGNGGGAGGCTTGAAAGGGTGTGAGCAGCACTGAT 600
QY 472 IleGlyGlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArg 491
Db 601 ATTGGTGGAGATGGCAGCTTGCAATGGAAATGGCAN-----AGANGGAAGTGAGGATGG 654
QY 492 LysGluGlyGlyPheLysArgValTyrLeuHisGlnGlu 504
Db 655 TAAGAAGGAAGTGGTTCAAAGAGTTCTTACACCAAGAG 693

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2004, 11:33:18 ; Search time 89.8744 Seconds
(without alignments)
5121.591 Million cell updates/sec

Title: US-10-051-909-32
Perfect score: 4075
Sequence: 1 IRSGSWLAVQTPTDLDLR.....PLEVITEFFAVGAQAAAKA 800

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_02:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3397	83.4	746	2 Q6H536	Q6h536 oryza sativ
2	3153	77.4	745	2 Q7X9Y6	Q7x9y6 saccharum h
3	2834	69.5	740	2 Q7XCM9	Q7xcm9 oryza sativ
4	2834	69.5	740	2 Q9FRK0	Q9frk0 oryza sativ
5	2808	68.9	743	2 Q8GT52	Q8gt52 hordeum vul
6	2654	65.1	739	2 Q8LPQ8	Q8lpq8 arabidopsis
7	2584	63.4	729	2 Q65497	Q65497 arabidopsis
8	2278	55.9	734	2 Q9SYQ3	Q9syq3 arabidopsis
9	2275	55.8	734	2 Q9LM67	Q9lm67 arabidopsis
10	2245	55.1	729	2 Q9SD00	Q9sd00 arabidopsis
11	2191.5	53.8	753	2 Q8GT51	Q8gt51 hordeum vul
12	2183	53.6	734	2 Q96290	Q96290 arabidopsis
13	1963	48.2	652	2 Q6K967	Q6k967 oryza sativ
14	1763	43.3	714	2 Q8H887	Q8h887 oryza sativ
15	1763	43.3	723	2 Q9SNK7	Q9snk7 oryza sativ
16	1270.5	31.2	746	2 Q8LJX7	Q8ljx7 sorghum bic
17	1050	25.8	643	2 Q8H6J2	Q8h6j2 zea mays (m
18	621.5	15.3	664	2 Q9XEV1	Q9xev1 oryza sativ
19	587.5	14.4	457	2 P96742	P96742 bacillus su
20	542	13.3	506	2 Q7XUE4	Q7xue4 oryza sativ
21	536	13.2	580	2 Q9C757	Q9c757 arabidopsis
22	535.5	13.1	581	2 Q9LKH2	Q9lkh2 mesembryant
23	533	13.1	461	1 CSBC BACSU	P46333 bacillus su
24	527.5	12.9	498	2 Q84UY4	Q84uy4 mesembryant
25	524.5	12.9	580	2 Q9ZQP6	Q9zqp6 arabidopsis
26	520.5	12.8	582	2 Q23492	Q23492 arabidopsis
27	519	12.7	581	2 Q9LKH1	Q9lkh1 mesembryant
28	518.5	12.7	469	2 Q97JE7	Q97je7 clostridium
29	512.5	12.6	577	2 Q945E5	Q945e5 oryza sativ
30	512	12.6	511	2 Q6ZFM6	Q6zfm6 oryza sativ
31	512	12.6	511	2 BAC83311	Bac83311 oryza sat

ALIGNMENTS

RESULT 1

ID	Q6H536	PRELIMINARY;	PRT;	746 AA.
AC	Q6H536;			
DT	05-JUL-2004 (Tremblrel. 27, Created)			
DT	05-JUL-2004 (Tremblrel. 27, last sequence update)			
DT	05-JUL-2004 (Tremblrel. 27, Last annotation update)			
DE	Putative hexose transporter.			
GN	Name=OSJNBb0035N08.17;			
OS	Oryza sativa (japonica cultivar-group).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;			
OC	Ehrhartoideae; Oryzeae; Oryza.			
OX	NCBI_TaxID=39947;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Sasaki T., Matsumoto T., Katayose Y.;			
RL	Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.			
CC	-!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).			
CC	-!- SIMILARITY: Belongs to the sugar transporter family.			
DR	EMBL; AP005756; BAD26163.1; -.			
DR	InterPro; IPR007114; MFS.			
DR	InterPro; IPR000508; Peptidase_S26.			
DR	InterPro; IPR005828; Sub_transporter.			
DR	InterPro; IPR003663; Sugar_transpt.			
DR	InterPro; IPR005829; Sug_transporter.			
DR	Pfam; PF00083; Sugar_tr; 1.			
DR	PRINTS; PR00171; SUGTRNSPORT.			
DR	PROSITE; PS00850; MFS; 1.			
DR	PROSITE; PS00501; SPASE_1; UNKNOWN_1.			
DR	PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.			
DR	PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.			
KW	Sugar transport; Transmembrane; Transport.			
SQ	SEQUENCE 746 AA; 80108 MW; 61F5890E2BFED02E CRC64;			

Query Match 83.4%; Score 3397; DB 2; Length 746;
Best Local Similarity 88.2%; Pred. No. 2.5e-212;
Matches 655; Conservative 40; Mismatches 48; Indels 0; Gaps 0;

Qy	56	MSGAVLVAIVASIGNLLQGWDNATIAAAVLYIKKEFQIQNEPTVEGLIVMSLIGATIVT	115
Db	1	MSGAAALVAIAASIGNLLQGWDNATIAAGAVLYIKKEFKLESEPTVEGLIVAMSLIGATIIIT	60
Qy	116	TFSGPLSDSIGRRPMLILSSILYFFSGGLIMLSPNVYVLLLARFVDFGIGLAVTLVPLY	175
Db	61	TFSGPVSVDWIGRRPMLILSSILYFLSSLIMLSPNVYVLLLARLIDGFGIGLAVTLVPLY	120
Qy	176	ISETAPSEIRGLLNTLPQFSGSGGMFLSYCMVFGMSLSPSPDWRIMLVLAIPSLFFFGL	235
Db	121	ISETAPSEIRGLLNTLPQFSGSGGMFLSYCMVFGMSLSPSPDWRIMLVLAIPSLFFFGL	180
Qy	236	TIFYLPESPRWLVSKGRMAEAKKVLQKLRGKDDVSGSELSLLLEGLEVGGTSTIEEYIIGP	295
Db	181	TIFYLPESPRWLVSKGRMAEAKKVLQKLRGREDVSGEMALLVEGLEVGADTSTIEEYIIGP	240

KW Sugar transport; Transmembrane.
SQ SEQUENCE 745 AA; 79955 MW; 54B12A7BFBDB8B155 CRC64;
Query Match 77.4%; Score 3153; DB 2; Length 745;
Best Local Similarity 80.7%; Pred. No. 1.9e-196;
Matches 601; Conservative 67; Mismatches 77; Indels 0; Gaps 0;
QY 296 ATEAADLVTGDKEQITLYGPEGQSWIARPSKGPIMLGSVLSASRHGSMVNSQSVPLM 355
Db 241 AIEPADEHVVDGDKQITLYGPEGQSWIARPSKGPISILGSLVSLTSRHGSMVNSQSVPLM 300
QY 356 DPIVTLFGSVHENMPQAGGSMRSTLFPNFGSMFVTDQHKNEQWDEENLHRDDDEYASD 415
Db 301 DPIVTLFGSVHENMPHAGGSMRSTLFPNFGSMFVTDQHPKVDQWDEENLHRDDDEYASD 360
QY 416 GAGGDYEDNLHSPLLSRQATGAFGKDIVHHGHRGSALSMMRQTLLGEGGDGVSTDDIGGG 475
Db 361 GAGGDYEDNVHSPLLSRQTTSAFGKDIVHHGHRGSALSMMRRTSLLEEGGEAVSTGIGGG 420
QY 476 WQLAWKSEKEGEGNGRKEGFKRIVYLHQEGVPGSRGSIYSLPGGGDVFECSFVHAAAL 535
Db 421 WQLAWKWSEREGEDGKKEGFKRIYVYLHQEEVPGSRGSIYSLPGGGDAPEGSEFIHAAAL 480
QY 536 VSQSALFSGKLAEPMSDAMVHPSEVAAGSRWKDLFEPGVRRLALLVGVGIQILQQFAG 595
Db 481 VSQALYSKDIIEQRMGPMTHPSEAAAGSSWKDLFEPGVRRLALLVGVGIQILQQFAG 540
QY 596 INGVLVYTPQILEQAGVAVILSKFGLSSASASILISLTLMLPCIGFAMLLMDLSGRR 655
Db 541 INGVLVYTPQILEQAGVAVILSNLGLSSASASILISLTLMLPSIGLAMRLMDISGRR 600
QY 656 FLLLGTIPILIASLVILVNSNIDLTGLTAHALLSTVSVIVYFCCFVMGFGPIPNILCAEI 715
Db 601 FLLLGTIPVLIASLVILVNSNIDLTGVAHAAALSTISVIYFCCFVMGFGPIPNILCAEI 660
QY 716 FPTRVRGLCIACAFTFWIGDIIVTYSLPVMNLNAIGLAGVFSIYAVVCLISFVFLKVP 775
Db 661 FPTRVRGICIAICALTFWIGDIIVTYSLPVMNLNAIGLAGVFSIYAVVCSIAFVFLKVP 720
QY 776 ETKGMPLEVITEFFAVGAKQAAA 798
Db 721 ETKGMPLEVITEFFAVGAKQMQA 743

RESULT 2
Q7X9Y6 PRELIMINARY; PRT; 745 AA.
AC Q7X9Y6;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative sugar transporter type 2a.
OS Saccharum hybrid cultivar.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Saccharum;
OC Saccharum officinarum complex.
OX NCBI_TaxID=128810;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22740598; PubMed=12856943;
RA Casu R.E., Grof C.P.L., Rae A.L., McIntyre C.L., Dimmock C.M.,
RA Manners J.M.;
RT "Identification of a novel sugar transporter homologue strongly
RT expressed in maturing stem vascular tissues of sugarcane by expressed
RT sequence tag and microarray analysis."
RL Plant Mol. Biol. 52:371-386(2003).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
DR EMBL: AY165599; AAO37640.1; -;
DR GO: GO:0016021; C: integral to membrane; IEA.
DR GO: GO:0005351; F: sugar porter activity; IEA.
DR GO: GO:0005215; F: transporter activity; IEA.
DR GO: GO:0006810; P: transport; IEA.
DR InterPro: IPR007114; MFS.
DR InterPro: IPR005828; Sub_transporter.
DR InterPro: IPR005829; Sug_transporter.
DR Pfam: PF00083; Sugar tr; 1.
DR PROSITE: PS00850; MFS; 1.
DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.

KW Sugar transport; Transmembrane.
SQ SEQUENCE 745 AA; 79955 MW; 54B12A7BFBDB8B155 CRC64;
Query Match 77.4%; Score 3153; DB 2; Length 745;
Best Local Similarity 80.7%; Pred. No. 1.9e-196;
Matches 601; Conservative 67; Mismatches 77; Indels 0; Gaps 0;
QY 56 MSGAVLVAIVASIGNLLQWGNATIAAAVLYIKKEFQIQNEPTVEGLIVMSLIGATIVT 115
Db 1 MSGAALVAIVASIGNLLQWGNATISAALLYIKKEFKLESEPTVEGLIVAMSLICATIT 60
QY 116 TFSGPLSDSIGRRPMLILSIYFFSGLIMLWSPNVYVLLARFVDGFGIGLAVTLVPLY 175
Db 61 TFSGPVSDWIGRLPMLILSVLYFVSSLIMLWSPNVYVLLARLINGFGVGLAVTLVPLY 120
QY 176 ISEIAPSEIRGLLNTLPQFSGSGMFLSYCMVFGMSLSPSPDWRIMLGVAIPSLFFFG 235
Db 121 ISETAPPEIRGLLNTLPQFSGSGMFLSYCMVFGMSLSPSPDWRIMLGVAIPSLFFFG 180
QY 236 TIFYPESPRWLVS KGRMAEAKKVLQRLRGKDDVTGEMALLVEGLEVGDDTSIEEYIIGP 295
Db 181 TIFYPESPRWLVS KGRMAEAKKVLQRLRGKDDVTGEMALLVEGLEVGDDTFIEEYIIGP 240
QY 296 ATEAADLVTGDKEQITLYGPEGQSWIARPSKGPIMLGSVLSASRHGSMVNSQSVPLM 355
Db 241 AIEPADHDVADGRERIILFPEGQSWIARPSKGSMLGSLVSGRHGSMVNSQSVPLM 300
QY 356 DPIVTLFGSVHENMPQAGGSMRSTLFPNFGSMFVTDQHKNEQWDEENLHRDDDEYASD 415
Db 301 DPIVTLFGSVHENISLAGGSMRIRVPSKFDMSISITDQHPKTEQWDEENLHRDDDEYATD 360
QY 416 GAGGDYEDNLHSPLLSRQATGAEKDIVHHGHRGSALSMMRQTLLGEGGDGVSTDDIGGG 475
Db 361 GAADYENNVHSPLLSRQNTSAEGKIAHGHGHRGSALGRRSLSDGEGEATSTGIGGG 420
QY 476 WQLAWKSEKEGEGNGRKEGFKRIVYLHQEGVPGSRGSIYSLPGGGDVFECSFVHAAAL 535
Db 421 WQLAWKWSEREGEDGKKEGFKRIYVYLHQEGVAGSRGSIYSLPGGGDAPEGSEFIHAAAL 480
QY 536 VSQSALFSGKLAEPMSDAMVHPSEVAAGSRWKDLFEPGVRRLALLVGVGIQILQQFAG 595
Db 481 VSQALYPRDITEQRMAGPATMHPSEAAAKVPSWRDLFEPGVRRLALLVGVGIQILQQFAG 540
QY 596 INGVLVYTPQILEQAGVAVILSKFGLSSASASILISLTLMLPCIGFAMLLMDLSGRR 655
Db 541 INGVLVYTPQIMEQAGVAVILSNLGLSSASASILISLTLMLPSIGLAMRLMDVSGRR 600
QY 656 FLLLGTIPILIASLVILVNSNIDLTGLTAHALLSTVSVIVYFCCFVMGFGPIPNILCAEI 715
Db 601 FLLLSTIPVLIASLVILVNSNIELGTVAHVLSTISVITLCCFKMGFGPIPNILCAEF 660
QY 716 FPTRVRGLCIACAFTFWIGDIIVTYSLPVMNLNAIGLAGVFSIYAVVCLISFVFLKVP 775
Db 661 FPTRVRGICIAICALIFWVGDIIVTYSLPVMNLNAIGLEGVFGIYAVACAIAFVFLKVP 720
QY 776 ETKGMPLEVITEFFAVGAKQAAA 800
Db 721 ETKGMPLEVITEFFAVGAKQAVAKA 745

RESULT 3
Q7XCM9 PRELIMINARY; PRT; 740 AA.
ID Q7XCM9
AC Q7XCM9;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative sugar transporter.
GN ORFNames=OSJNBb0064P21.3;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartioideae; Oryzeae; Oryza.

OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA The Rice Chromosome 10 Sequencing Consortium;
RT "In-depth view of structure, activity, and evolution of rice
RT chromosome 10."
RL Science 300:1566-1569(2003).
RN [2]
RP SEQUENCE FROM N.A.
RA Buell C.R., Wing R.A., McCombie W.R., Messing J., Yuan Q.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
DR EMBL; AE017116; AAP54842.1; --
DR Gramene; Q7XCM9; --
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005351; F:sugar porter activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF00083; Sugar_tr; 1.
DR PROSITE; PS00850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Sugar transport; Transmembrane.
SQ SEQUENCE 740 AA; 78898 MW; D4C003E0A4980C46 CRC64;

Query Match 69.5%; Score 2834; DB 2; Length 740;
Best Local Similarity 74.5%; Pred. No. 1e-175;
Matches 553; Conservative 75; Mismatches 108; Indels 6; Gaps 5;

QY 56 MSGAVLVAIVASIGNLLQGDNATIAAAVLYIKKEFQIQNEPTVEGLIVMSLIGATIVT 115
DB 1 MAGAVLVAIAASIGNLLQGDNATIAAGAVLYIKKEFNQSEPLIEGLIVAMSLIGATIIT 60

QY 116 TFSGPLSDSIGRRPMLILSSILYFFSGLIMLWSPNVYVLLARFVDGFGIGLAVTLVPLY 175
DB 61 TFSGAVADSFGRRPMLIASAVLYFVSGVLWMLWAPNVYVLLARLIDGFGIGLAVTLVPLY 120

QY 176 ISEIAPSEIRGLNLTLPQFSGSGGMFLSYCMVFGMSLSPSPDWRIMLGVLAIPLSFFFG 235
DB 121 ISETAPTDIRGLNLTLPQFSGSGGMFLSYCMVFGMSLMPQPDWRIMLGVLISPLIYFAL 180

QY 236 TIFYLPESPRWLVSKGRMAEAKVLOKLRGDDVSGELSLLEGLVGGDTISIEEYIIGP 295
DB 181 TIFYLPESPRWLVSKGRMAEAKRVLQGLRGEDVSGEMALLVEGLGVGKDKIEEYIIGP 240

QY 296 ATEAADLVTGDKEQITLYGPEEGQSWIARPSKGPIMGLSVLSASRHGSMVNSQVPLM 355
DB 241 DDELADEGLAP-DPEKIKLYGPEEGLSWVARPVHGQSALGSALGLISRHGSMVNSQKPLV 299

QY 356 DPVITLFGSVHNNPQAGGSMRSTLFPNFGMFSVTDQHAQNEQWDEENLHRDDEEYASD 415
DB 300 DPVITLFGSVHNNPQAGGSMRSTLFPNFGMFSVAEQQAQKGDWDAES-QREGEDYGS 358

QY 416 GAGGDYEDNLHSPILSRQATGAEGKDI-VHHGRGSALSMMRQTLLEGGDGVSTDIGG 474
DB 359 HGGDDIEDLSQSPILSRQATSVEGKEIAAPHGSIINGAVG--RSSLMQGEAVSSMGIGG 416

QY 475 GWQLAWKWEKEGNGRKEGGFKRVYLHQGVGPGRRGSIIVSLPGGDFVEGSEFVHAAA 534
DB 417 GWQLAWKWTREGEADGEGGFKRIYLYHEEGVTGDRRGSIIVSLP-GGDVPPGGEFVQAAA 475

QY 535 LVQSALFSKGLAEPRMSDAAMVHPSEVAAGSRKWKDLFEPGVRRALLVGVGIQILQOQFA 594
DB 476 LVQPALYSKELMEORLAGPAMVHPSPQAVAKGPKWADLFEPGVKHALFVGIGIQLQOQFA 535

QY 595 GINGVLYYTPQILEQAGVAVILSKFGLSSASASILISLTLLMLPCICFAMLLMDLSGR 654
DB 536 GINGVLYYTPQILEQAGVGVLLANIGLSSSSASILISGLTLLMLPSIGIAMRLMDMSGR 595

QY 655 RFLLLGTIPILIASLVILVSNLIDGLTAHALLSTVSIVYFCCFVGMGFGPIPNILCAE 714

DB 596 RFLLLATIPILIALILVILNLDVGTWVTRASLSTVSVILYFCFFVGMGFGPIPNILCAE 655
QY 715 IFPTRVRGLCIAICAFTEWIDIIIVTYSPLVMLNAIGLAGVFSIYAVVCLISFVFLKV 774
DB 656 IFPTTVRGICIAICAFTEWIDIIIVTYSPLVMLNAIGLAGVFSIYAVVCLISFVFLKV 715
QY 775 PETKGMPLVITEFEFAVGAQQA 796
DB 716 PETKGMPLVITEFEFAVGAQQA 737

RESULT 4
Q9FRKO PRELIMINARY; PRT; 740 AA.
AC Q9FRKO;
DT 01-MAR-2001 (TremBLrel. 16, Created)
DT 01-MAR-2001 (TremBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Putative sugar transporter.
GN Name=OSJNBb0064P21.3;
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Buell C.R., Yuan Q., Moffat K.S., Hill J.N., Burr P.C., Hsiao J.,
RA Zismann V., Pai G., Bowman C.L., Fujii C.Y., VanAken S.E.,
RA Bowman C.L., Craven B., Utterback T.R., Khalak H., Feldblyum T.V.,
RA Quackenbush J., White O., Salzberg S.L., Fraser C.M.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Buell R.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the sugar transporter family.
DR EMBL; AC073166; AAG46115.1; --
DR Gramene; Q9FRKO; --
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005351; F:sugar porter activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006843; P:carbohydrate transport; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR003663; Sug_transport.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF00083; Sugar_tr; 1.
DR PROSITE; PRO0171; SUGTRNSPORT.
DR PROSITE; PS00850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Sugar transport; Transmembrane; Transport.
SQ SEQUENCE 740 AA; 78898 MW; D4C003E0A4980C46 CRC64;

Query Match 69.5%; Score 2834; DB 2; Length 740;
Best Local Similarity 74.5%; Pred. No. 1e-175;
Matches 553; Conservative 75; Mismatches 108; Indels 6; Gaps 5;

QY 56 MSGAVLVAIVASIGNLLQGDNATIAAAVLYIKKEFQIQNEPTVEGLIVMSLIGATIVT 115
DB 1 MAGAVLVAIAASIGNLLQGDNATIAAGAVLYIKKEFNQSEPLIEGLIVAMSLIGATIIT 60

QY 116 TFSGPLSDSIGRRPMLILSSILYFFSGLIMLWSPNVYVLLARFVDGFGIGLAVTLVPLY 175
DB 61 TFSGAVADSFGRRPMLIASAVLYFVSGVLWMLWAPNVYVLLARLIDGFGIGLAVTLVPLY 120

QY 176 ISEIAPSEIRGLNLTLPQFSGSGGMFLSYCMVFGMSLSPSPDWRIMLGVLAIPLSFFFG 235
DB 121 ISETAPTDIRGLNLTLPQFSGSGGMFLSYCMVFGMSLMPQPDWRIMLGVLISPLIYFAL 180

QY 236 TIFYLPESPRWLVSKGRMAEAKVLOKLRGDDVSGELSLLEGLVGGDTISIEEYIIGP 295
DB 181 TIFYLPESPRWLVSKGRMAEAKRVLQGLRGEDVSGEMALLVEGLGVGKDKIEEYIIGP 240

QY 296 ATEAADLVTGDKEQITLYGPEEGQSWIARPSKGPIMGLSVLSASRHGSMVNSQVPLM 355
DB 241 DDELADEGLAP-DPEKIKLYGPEEGLSWVARPVHGQSALGSALGLISRHGSMVNSQKPLV 299

QY 356 DPVITLFGSVHNNPQAGGSMRSTLFPNFGMFSVTDQHAQNEQWDEENLHRDDEEYASD 415
DB 300 DPVITLFGSVHNNPQAGGSMRSTLFPNFGMFSVAEQQAQKGDWDAES-QREGEDYGS 358

QY 416 GAGGDYEDNLHSPILSRQATGAEGKDI-VHHGRGSALSMMRQTLLEGGDGVSTDIGG 474
DB 359 HGGDDIEDLSQSPILSRQATSVEGKEIAAPHGSIINGAVG--RSSLMQGEAVSSMGIGG 416

QY 475 GWQLAWKWEKEGNGRKEGGFKRVYLHQGVGPGRRGSIIVSLPGGDFVEGSEFVHAAA 534
DB 417 GWQLAWKWTREGEADGEGGFKRIYLYHEEGVTGDRRGSIIVSLP-GGDVPPGGEFVQAAA 475

QY 535 LVQSALFSKGLAEPRMSDAAMVHPSEVAAGSRKWKDLFEPGVRRALLVGVGIQILQOQFA 594
DB 476 LVQPALYSKELMEORLAGPAMVHPSPQAVAKGPKWADLFEPGVKHALFVGIGIQLQOQFA 535

QY 595 GINGVLYYTPQILEQAGVAVILSKFGLSSASASILISLTLLMLPCICFAMLLMDLSGR 654
DB 536 GINGVLYYTPQILEQAGVGVLLANIGLSSSSASILISGLTLLMLPSIGIAMRLMDMSGR 595

QY 655 RFLLLGTIPILIASLVILVSNLIDGLTAHALLSTVSIVYFCCFVGMGFGPIPNILCAE 714

QY 236 TIFYPESPRWLVSKGRMAEAKVQLKRGKDDVSGELSLLEGLVGGDTSIEEYIIGP 295
Dp 181 TIFYPESPRWLVSKGRMAEAKVQLKRGREDVSGEMALLVEGLVGKDTKIEEYIIGP 240
QY 296 ATEAADLVTDGDKQITLYGPEEGQSWIARPSKGPIMLGSVLSASRHGSMVNSQVPLM 355
Dp 241 DDELADEGLAP-DPEKIKLYGPEEGLSWVARPVHGQSALGSLISRHGSMVNSQVPLV 299
QY 356 DPIVTLFGSVHENMPQAGGMRSTLFPNFGSMFSTVDQAHKNEQWDEENLHRDDEEYASD 415
Dp 300 DPVVTFLFGSVHEKMPPEIMSGMRSTLFPNFGSMFSAEQQAQKGDWDAES-QREGEDYGS 358
QY 416 GAGGDYEDNLHSPLLSRQATGAEGKDI-VHHGHRGSALSMMRRQTLLGEGDGVSTDIGG 474
Dp 359 HGGDDIEDSLQSPILSRQATSVGEKETAAPHGSIIMGAVG--RSSSLMQGGEAVSSMGIGG 416
QY 475 GWQLAWKWKSEKENGKRGKGFYVYLHQEGVPGSRGSGIVSLPGGGDVFESEFVHAAA 534
Dp 417 GWQLAWKWKTEREGADGEKEGGFQRIYLYHEEGVTGDRRGSILSLP-GGDVPPGGEFVQAAA 475
QY 535 LVSQSALFSKGLAEPRMSDAAMVHPSEVAAGSRWKDLFEPGVRRALLVGVGIQILQQA 594
Dp 476 LVSQPALYSKELMEQRLAGPAMVHPSPQAVAKPKWADLFEFGVKHALFVGIGIQLQQA 535
QY 595 GINGVLYYTPQILEQAGVAVILSKFGLSSASASILISSLTLLMLPCIGFAMLLMDLSGR 654
Dp 536 GINGVLYYTPQILEQAGVGVLLANIGLSSSSASILISGLTLLMLPSIGIAMRLMDMSGR 595
QY 655 RFLLLGTIPILIASLVILVSNLIDLGTLAHALLSTVSIVYFCCFVMGFGPIPNILCAE 714
Dp 596 RFLLLATIPILIVALAILVILNILDVGTVMHASLSTVSIVLYFCFFVMGFGPIPNILCAE 655
QY 715 IFPTVRVGLCIAICAFTEWIGDIIIVTSLPVMNLNAIGLAGVFSIYAVVCLISFVFLKV 774
Dp 656 IFPTVRVGLCIAICALTFWIGDIIIVTSLPVMNLNAIGLAGVFSIYAVVCLILAFVFMKV 715
QY 775 PETKGMPLLEVITEFFAVGAKQA 796
Dp 716 PETKGMPLLEVITEFFSVGAKQA 737

RESULT 5
Q8GT52 PRELIMINARY; PRT; 743 AA.
ID Q8GT52;
AC Q8GT52;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Hexose transporter.
GN Name=stpl;
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
OC Triticeae; Hordeum.
OX NCBI_TaxID=4513;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Developing caryopsis;
RX MEDLINE=22424051; PubMed=12535352;
RA Weschke W., Panitz R., Gubatz S., Wang Q., Radchuk R., Weber H.,
RA Wobus U.;
RT "The role of invertases and hexose transporters in controlling sugar
RT ratios in maternal and filial tissues of barley caryopses during early
RT development.";
RL Plant J. 33:395-411(2003).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the sugar transporter family.
DR EMBL; AJ534445; CAD58958.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005351; F:sugar porter activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0008643; P:carbohydrate transport; IEA.
DR InterPro; IPR007114; MFS.

DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR003663; Sugar_transpt.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF00083; Sugar_tr; 1.
DR PRINTS; PR00171; SUGTRNSPORT.
DR PROSITE; PS00850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Sugar transporter; Transmembrane; Transport.
SQ SEQUENCE 743 AA; 79339 MW; 0169B6BD06007DF5 CRC64;
Query Match 68.9%; Score 2808; DB 2; Length 743;
Best Local Similarity 73.1%; Pred.No. 5.2e-174;
Matches 543; Conservative 84; Mismatches 104; Indels 12; Gaps 6;
QY 57 SGAVLVAIVASIGNLLQGWDNATIAAAVLYIKKEFQQLONEFTVEGLIVMSLIGATVIT 116
Db 3 SGAVLVAIAASIGNLLQGWDNATIAAGVLYIKKEFSLETQPLIEGLIVAMSLIGATVIT 62
QY 117 FSGPLSDSICRRPMLILSSILYFFESGLIMLWSPNVYVLLARFVDGFGIGLAVTLVPLY 176
Db 63 FSGAVADAVGRRRPLLIASSVLYFVSGLVMLWAPNVYVLLARLIDGFGIGLAVTLVPLY 122
QY 177 SEIAPSEIRGLLNTLPQFSGSGGMFLSYCMVFGMSLSPSPDWIRIMLGVLAIPSLFFFGLT 236
Db 123 SETAPTDIRGLLNTLPQFSGSGGMFLSYCMVFTMSLMPQPDWRIMLGVLISPSLMYFALT 182
QY 237 IFYLPESPRWLVSKGRMAEAKVQLKRGKDDVSGELSLLEGLVGGDTSIEEYIIGP 296
Db 183 VFYLPESPRWLVSKGRMAEAKVQLKRGREDVSGEMALLVEGLVGKDTHEEYIIGPD 242
QY 297 TEAADLVTDGDKQITLYGPEEGQSWIARPSK--GPIMLGSVLSASRHGSMVNSQVPL 354
Db 243 DELADDGLAP-DQEKLYGAEEGVSWIARPVXGGQSALGSALGMSRHGSMVNSQKSL 301
QY 355 MDPIVTLFGSVHENMPQAGGMRSTLFPNFGSMFSTVDQAHKNEQWDEENLHRDDEEYAS 414
Db 302 VDPLVTLFGSVHEKMPVMSGMRSTLFPNFGSMFSAEQQAQKADWDAES-HRDDDEYAS 360
QY 415 DGAGGDYEDNLHSPLLSRQATGAEGKDIV--HHGHRGSALSMMRRQTLLGEGDGVSTDI 472
Db 361 DHGADDIEDNLSPILSRQATSVGEKETAAPHGSIIMGVSSSMQ----GGDAVSSMGI 415
QY 473 GGGWQLAWKWKSEKENGKRGKGFYVYLHQEGVPGSRGSGIVSLPGGGDVFESEFVHA 532
Db 416 GGGWQLAWKWKTEREGADGQKEGFGRIYLYHEEGVSGDRGSIILSMP-GGDIPPGGEYIQA 474
QY 533 AALVSQSALFSKGLAEPRMSDAAMVHPSEVAAGSRWKDLFEPGVRRALLVGVGIQILQQA 592
Db 475 AALVSQPALYSKDLIEQQLAGPAMVHPSEAVAKGKWAELFEPGVKHALFVGIGLQILQQ 534
QY 593 FAGINGVLYYTPQILEQAGVAVILSKFGLSSASASILISSLTLLMLPCIGFAMLLMDLS 652
Db 535 FAGINGVLYYTPQILEQAGVGVLLSNIGLSSSSASILISALTLLMLPSIGIAMRLMDMS 594
QY 653 GRRFLLGTIPILIASLVILVSNLIDLGTLAHALLSTVSIVYFCCFVMGFGPIPNILC 712
Db 595 GRRFLLSTIPVLIVAILVNLVLDVGTVMHAALSTISIVYFCFFVMGFGPIPNILC 654
QY 713 AEIPTVRVGLCIAICAFTEWIGDIIIVTSLPVMNLNAIGLAGVFSIYAVVCLISFVFL 772
Db 655 AEIPTSVRGICIAICALTFWIGDIIIVTSLPVMNLNAIGLAGVFSIYAVVCLIAFVFM 714
QY 773 KVPETKGMPLLEVITEFFAVGAKQ 795
Db 715 KVPETKGMPLLEVITEFFSVGAKQ 737

RESULT 6
Q8LPQ8 PRELIMINARY; PRT; 739 AA.
ID Q8LPQ8
AC Q8LPQ8;
DT 01-OCT-2002 (Tremblrel. 22, Created)

01-OCT-2002 (Tremblrel. 22, Last sequence update)
01-OCT-2003 (Tremblrel. 25, Last annotation update)
AT4g35300/F23E12.140.
Arabidopsis thaliana (Mouse-ear cress).
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
eurosids II; Brassicales; Brassicaceae; Arabidopsis.
NCBI_TaxID=3702;
[1]
SEQUENCE FROM N.A.
Rn Shinn P., Chen H., Cheuk R., Kim C.J., Meyers M.C., Banh J.,
Bowser L., Carninci P., Chang E., Dale J.M., Goldsmith A.D.,
Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,
Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,
Seki M., Southwick A., Tang C.C., Toriumi M., Wu H.C., Yamada K.,
Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
Ecker J.R.;
Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
-!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
-!- SIMILARITY: Belongs to the sugar transporter family.
EMBL; AY094465; AAM19835.1; -.
DR GO: 0016021; C: integral to membrane; IEA.
DR GO: 0005351; F: sugar porter activity; IEA.
DR GO: 0005215; F: transporter activity; IEA.
DR GO: 0008643; P: carbohydrate transport; IEA.
DR InterPro: IPR007114; MFS.
DR InterPro: IPR005828; Sub transporter.
DR InterPro: IPR003663; Sugar_transpt.
DR InterPro: IPR005829; Sug transporter.
DR Pfam; PF00083; Sugar tr; 1.
DR PRINTS; PR00171; SUGTRNSPT.
DR PROSITE; PS00850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Sugar transport; Transmembrane; Transport.
SQ SEQUENCE 739 AA; 79725 MW; 5D67C6D2DE7563FD CRC64;

Query Match 65.1%; Score 2654; DB 2; Length 739;
Best Local Similarity 70.0%; Pred. No. 5.4e-164;
Matches 521; Conservative 90; Mismatches 123; Indels 10; Gaps 7;

56 MSGAVLVAVIASTGNLLQGDNDATIAAAVLYIKKEFQIQNEPTVEGLIVSMISLIGATIVT 115
1 MSGAVLVAVIAAAVGNLLQGDNDATIAAGVLYIKKEFNLESNPSVEGLIVAMSLIGATLIT 60
116 TFSGLPLSDSIGRRPMLILSSILYFFSGLIMLWSPNVYVLLARFVDGFGIGLAVTLVPLY 175
61 TCSGGVADWLGRPRMLILSSILYFVGSVLMWSPNVYVLLGRLLDGFVGLVTLVPIY 120
176 ISEIAPSEIRGLNLTLPQFSGGGMFLSYCMVFGMSLSPSPDWRLMGLVLAIPSLFFGL 235
121 ISETAPSEIRGLNLTLPQFSGGGMFLSYCMVFGMSLSPSPDWRLMGLVLAIPSLFFFL 180
236 TIFYPESPRLVSKGRMAEAKKVLQKLRGKDDVSGELSLLEGLVGGDTSEIEYIIGP 295
181 TVFFLPESPRLVSKGRMLEAKRVLQRLRGREDVSGEMALLVEGLIGGGETTIEEYIIGP 240
296 ATEAADLVTDGDKQITLYGPEEGQSWIARPSKGPIMLGSVLSLRHGS-MVNQSVPL 354
241 ADEVTDDHDIADVDDQIKLYGAEGLSWVARPVKG---GSTMSVLSRHGSTMSTRQGS 296
355 MDPIVTLFGSVHENMPQAGGSMRSTLFPNFGSMFVSVDQAKNEQWDEENLHRDDEEYAS 414
297 IDPLVTLFGSVHEKMPDT-GSMRSALFPFHFGSMFVSQNPQPRHEDWDEENLVGEGEDYPS 355
415 DGAGGDYEDNLHSPLLSRQATGAEGKDIVHHGHRGSAISMRRTLLGEGGDSVSTDIGG 474
356 D-HGDDSEDDLHSPLLSRQTSME-KDMPHTAHGTLSTFRHGSQVQGAQGEAGSMGIGG 413
475 GWQLAWKWEKEGNGRKEGFKRVYLHQEGVPGSRGSIIVSLPGGGDVFESEFVHAA 534
414 GWQVANKWTEREDESQKEGFKRIYLHQEGFPGSRGSIIVSLPGGGDTGCE-ADFVQASA 472

QY 535 LVQSALFESKGLAEPRMSDAMVHPSEVAAGSRWKDLPEPGVRRALLVGVGIQILOQFA 594
Db 473 LVSQPALYSKDLLKEHTIGPAMVHPSE-TTKGSIWHLDPGVKRALVGVGLQILOQFS 531
QY 595 GINGVLYYTPQILEQAGVAVILSKFGLSSASASILISLTLLMLPCIGFAMLLMDLSGR 654
Db 532 GINGVLYYTPQILEQAGVIGLLSNMGISSSSASLLISALTTFVMLPAIAVAMRLMDLSGR 591
QY 655 RFLLLGTIPILIASLVILVSNLIDLTGLAHALLSTVSIVVFCECFVMGFGPIPNILCAE 714
Db 592 RTLLLTIPILIASLLVLVISNLVHMNSIVHAVLSTVSIVVFCECFVMGFGPAPNLCSE 651
QY 715 IFPTRVRGLCIACTFWIGDIIIVTSLPVMNLNAIGLGVFSIYAVVCLISFVFLKV 774
Db 652 IFPTRVRGICIACTFWICDIIIVTSLPVLKLSIGLGVFMVIAIVCCISWVFIKV 711
QY 775 PETKGMPLVITEFFAVGAKQAAA 798
Db 712 PETKGMPLVITEFFSVGARQAEA 735
RESULT 7
O65497
ID O65497 PRELIMINARY; PRT; 729 AA.
AC O65497;
DT 01-AUG-1998 (Tremblrel. 07, Created)
DT 01-AUG-1998 (Tremblrel. 07, Last sequence update)
DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)
DE Putative sugar transporter protein (Monosaccharide sensing protein
DE 2).
GN Name=F23E12.140; Synonyms=AT4g35300, mssp2;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Hilbert H., Braun M., Holzer E., Brandt A., Duesterhoeft A.,
RA Hoheisel J., Mewes H.W., Mayer K.F.X., Schueller C.;
RA Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Hilbert H., Braun M., Holzer E., Brandt A., Duesterhoeft A.,
RA Mewes H.W., Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA Stamme C., Tjaden J., Trentmann O., Emmerlich V., Neuhaus E.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the sugar transporter family.
DR EMBL; AL022604; CAA18739.1; -.
DR EMBL; AL161587; CAB80247.1; -.
DR EMBL; AJ532570; CAD58692.1; -.
DR PIR; T06127; T06127.
DR GO: 0016021; C: integral to membrane; IEA.
DR GO: 0005351; F: sugar porter activity; IEA.
DR GO: 0005215; F: transporter activity; IEA.
DR GO: 0008643; P: carbohydrate transport; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub transporter.
DR InterPro; IPR003663; Sugar_transpt.
DR InterPro; IPR005829; Sug transporter.
DR

DR Pfam; PF00083; Sugar tr; 1.
DR PRINTS; PR00171; SUGTRNSPORT.
DR PROSITE; PS50850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Sugar transport; Transmembrane; Transport.
SQ SEQUENCE 729 AA; 78524 MW; 81A35DDPF8725411E CRC64;

Query Match 63.4%; Score 2584; DB 2; Length 729;
Best Local Similarity 68.8%; Pred. No. 1.9e-159;
Matches 512; Conservative 89; Mismatches 123; Indels 20; Gaps 8;

QY 56 MSGAVLVAIVASIGNLLQGDWNTAAAVLYIKKEFQQLNEPTVEGLIVSMSLIGATIVT 115
Db 1 MSGAVLVAIAAAGVNLQGDWNTAAAGVLYIKKEFNLESNPSVEGLIVAMSLIGATLIT 60
QY 116 TFSGPLSDSIGRRPMLILSSILYFSGGLIMLWSPNVYVLLARFVDGFGIGLAVTLVPLY 175
Db 61 TCSGGVADWLGRPRMLILSSILYFVCSLVMLWSPNVYVLLGRLLDGFVGVLVTLVPIY 120
QY 176 ISEIAPSEIRGLLNTLPQFSGGGMFLSYCMVFGMSLSPSPDWIRIMLGVLAIPLSFFFL 235
Db 121 ISETAPPEIRGLLNTLPQFTSGGGMFLSYCMVFGMSLSPSPDWIRIMLGVLAIPLSFFFL 180
QY 236 TIFYPESPRWLVSCKRMAEAKVQLKRGKDDVSGELSLLEGLVGGDTSEIYIIGP 295
Db 181 TVFFLPSPRWLVSKGRMLEAKRVLQRLRGREDVSGEALLVEGLIGGETTIEYIIGP 240
QY 296 ATEAADLVTDGDKXEQITLYGPEEGQSWIARPSKGPIMLGSVLASRHGS-MVNQSVPL 354
Db 241 ADEVTDDHDIADVDDQIKLYGAEGLSWVAVPVKG---GSTMSVLSRHGSIMRRQGS 296
QY 355 MDPIVTLFGSVHENMPQAGGMRSTLFPNFGSMFSTVDQHAKEQWDEENLHRDDEYAS 414
Db 297 IDPLVTLFGSVHEKMPDT-GSMRSALFPFHFGSMFSGVGNQPRHEDWDEENLVEGEDYPS 355
QY 415 DGAGGYEDNLHSPLLSRQATGAEGKDIVHHGHRGSLASMRRTLLGEGGDVSGSTDIGG 474
Db 356 D-HGDDSEDDLHSPILSRQITSMV-KDMPHTAHGTLSTFRHGSQVQGAQEGAGSMGIGG 413
QY 475 GWQLAWKWEKEGEGNGRKEGKFRVYLHQEGVPGSRGSIIVSLPGGGDVFESEFVHAAA 534
Db 414 GWQVAKWTEREDESQKE-----EGFPGSRGSIIVSLPGGGDTGE-ADVFQASA 462
QY 535 LVQSALFSKGLAEPRMSDAAMVHPSEVAAKGSRWKDLFEPGVRRALLVGVGIQILQOFA 594
Db 463 LVSQPALYSKDLLKEHTTIGPAMVHPSE-TTKGSIWHDLDHPGVKRALVGVGLIQLQOFS 521
QY 595 GINGVLYYTPQILEQAGVAVILSKFGLSSASASILLSSLTLLMLPCIGFAMLLMDLSGR 654
Db 522 GINGVLYYTPQILEQAGVAVILSKFGLSSASASILLSSLTLLMLPAIAVAVRLMDLSGR 581
QY 655 RFLLLGTIPILIASLVLVSNLIDLGLTAHALLSTVSVIVYFCFVMGFGPIPNILCAE 714
Db 582 RLLLTTPILIASLVLVSNLHMNSIVHAVLSTVSVVLYFCFFVMGFGPAPNILCSE 641
QY 715 IFPTRVRGLCIAICAFTEWIGDIIIVTYSPLVNLNAIGLAVFSIYAVVCLISFVFLKV 774
Db 642 IFPTRVRGICIAICALTFWICDIIIVTYSPLVNLKISGLAGVFGMYAIVCCISWVFIKV 701
QY 775 PETKGMPLVITEFFAVGAKQAAA 798
Db 702 PETKGMPLVITEFFSVGARQAEA 725

RESULT 8

Q9SYQ3 ID Q9SYQ3 PRELIMINARY; PRT; 734 AA.
AC Q9SYQ3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Sugar transporter.

GN Name=PF00083; Sugar tr; 1.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsais.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Federspiel N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,
RA Altafi H., Araujo R., Huizar L., Rowley D., Buehler E., Dunn P.,
RA Gonzalez A., Kremenetskaia I., Kim C., Lenz C., Li J., Liu S.,
RA Lueros S., Schwartz J., Shinn P., Toriumi M., Vysotskaia V.S.,
RA Walker M., Yu G., Ecker J., Theologis A., Davis R.W.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
CC -|- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -|- SIMILARITY: Belongs to the sugar transporter family.
DR EMBL; AC007369; AAD30608.1; -.
DR PIR; H86340; H86340.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005351; F:sugar porter activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0008643; P:carbohydrate transport; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR003663; Sugar_transpt.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF00083; Sugar tr; 1.
DR PRINTS; PR00171; SUGTRNSPORT.
DR PROSITE; PS50850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Sugar transport; Transmembrane; Transport.
SQ SEQUENCE 734 AA; 79425 MW; AB027864F850259C CRC64;

Query Match 55.9%; Score 2278; DB 2; Length 734;
Best Local Similarity 61.2%; Pred. No. 1.6e-139;
Matches 462; Conservative 112; Mismatches 145; Indels 36; Gaps 14;

QY 56 MSGAVLVAIVASIGNLLQGDWNTAAAVLYIKKEFQQLNEPT-VEGLIVSMSLIGATIV 114
Db 1 MKGATLVALAATIGNFLQGDWNTAAGAMVYINKDLNL---PTSVQGLVVAMSLIGATVI 57
QY 115 TFSGPLSDSIGRRPMLILSSILYFSGGLIMLWSPNVYVLLARFVDGFGIGLAVTLVPL 174
Db 58 TCSGGPISDWLGRPRMLILSSVYFVCGGLIMLWSPNVYVLCFARLLNGFGAGLAVTLVPV 117
QY 175 YISEIAPSEIRGLLNTLPQFSGGGMFLSYCMVFGMSLSPSPDWIRIMLGVLAIPLSFFFG 234
Db 118 YISETAPPEIRGLLNTLPQFSGGGMFLSYCMVFTMSLSDSPSWRAMLGVLISPLLYLF 177
QY 235 LTIFYLPESPRWLVSCKRMAEAKVQLKRGKDDVSGELSLLEGLVGGDTSEIYIIG 294
Db 178 LTVFYLPEPRWLVSCKRMAEAKRVLQQLCGREDVTKMALLVEGLDIGEKTMEDLLVT 237
QY 295 PATEAADLVTDGDKXEQITLYGPEEGQSWIARPSKGPIMLGSVLASRHGSIMVNQSV 353
Db 238 LEDHEGDDTLETVDGQMRLYGTHENQSYLARVPVPEQ---NSSGLFSRHGSLANQSMI 294
QY 354 LMDPIVTLFGSVHENMPQAGGMRSTLFPNFGSMFSTVDQ--HAKNEQWD---EENLHRD 408
Db 295 LKDPVLNLFGLSHEKMEPEAGGNTRSGIFPHFGSMFSTTADAPHGKPAHWEKDIESHYKD 354
QY 409 DEEYASDGAGGDYED--NLHSPLLSRQATGAEGKDIVHHGHRGSLASMRRTLLGEGGD 465
Db 355 NDDYATDDGAGDDDDSDNLSPLMSRQTTSMD-KDMI PHPTSGSTLSMRHSTLMQ-GN 412
QY 466 GVSSTDIGGWLAWKWEKEGEGNGRKEGKFRVYLHQEGVPGSRGSIIVSLPGGGDVFE 525
Db 413 GESSMGIGGGWHMGYRYENDE-----YKRYLYKEDGAE-SRSGIISIPGGPD--G 460
QY 526 GSEFVHAAALVSQALFSKGLAEPRMSDAAMVHPSEVAAKGSRWKDLFEPGVRRALLVGV 585
Db 461 GGSYIHASALVSRSLGPKS-----VHGSAMVPEKIAASGPLWSALLEPGVKRALVGV 515

Qy	586	GIQILQQFAGINGVLYYTPQILEQAGVAVILSKFGLSSASASILSSLTTLMLPCIGFA	645
Db	516	GIQILQQFSGINGVLYYTPQILERAGVDILLSSGLSSISASFLISGLTTLMLPAIVVA	575
Qy	646	MLLMDLSGRRFLLLTIPILLIASLVILVVSNLIDLGLTAHALLSTVSVIVYFCFVMGFG	705
Db	576	MRLMDVSGRRSLLLTIPVLIVSLVVLVISELIHISKVYNAALSTGCVLYFCFFVMGYG	635
Qy	706	PIPNILCAEIFFPTRVRGLCIAICAFTEFWIGDIIIVTYSLPVMLNAIGLGVFSIYAVVCLI	765
Db	636	PIPNILCSEIFFPTRVRGLCIAICAMVFWIGDIIIVTYSLPVLLSSIGLVGFSIYAAVCVI	695
Qy	766	SFVFVFLKVPETKGMPLFVITEFFFAVGAQKQAAKA	800
Db	696	SWIFVMKVPETKGMPLFVITDYFAFGA-QAQASA	729

RESULT	9
Q9LM67	
ID	Q9LM67
AC	Q9LM67; PRELIMINARY; PRT; 734 AA.
DT	01-OCT-2000 (TrEMBLrel. 15, Created)
DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE	F2D10.36.
OS	Arabidopsis thaliana (Mouse-ear cress).
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsi.
OX	NCBI_TaxID=3702;
FN	[1]

Qy	56	MSGAVLVAIVASIGNLLQGWDNATIAAAVLVIKKEFQLONEPT-VEGLIVSMSLGATIV	114
Db	1	MKGATLVALAATIGNFLOGWDNATIAGAMVYINKDLNL---PTSVOGLVVAMSLGATIVI	57
Qy	115	TTCSGPLSDSIGRRPMLILSSILYFFSGLIMLWSPNVYVLLARFVDGFGIGLAIVTLVPL	174
Db	58	TTCSGPISDLWRPRMLILSSVMYFVCGLIMLWSPNVYVLCFARLLNGFGAGLAIVTLVPV	117
Qy	175	YISEIAPSEIRGLLNTLPQFSGGGMFLSYCMVFMSLSPPSPDWRIMLGVLAIPSLFFFEG	234
Db	118	YISETAPPEIRGQLNTLPQFLGSGGMFLSYCMVFTMSLSDSPSWRAMLGVLSIPSLLYLF	177
Qy	235	LTI FYLPESPRWLVS KGRMAEAKVKLQKLRGKDDVSGELSLLEGLEVGGDTSIEEYIIIG	294
Db	178	LTVFYLPESPRWLVS KGRMDEAKRVLQQCGREDVTDEMALLVEGLDIGGEKTMEDLLVT	237
Qy	295	PATEAADLVTDGDKE-QITLYGPEEQSQWTARPSPKGPIMLGSVLSLASRHGSMVNQSV	353
Db	238	LEDHEGDTLETVDGEDQMRLYGTHENQSYLARPVEQ---NSSLGLRSRHRHGLANQSMI	294
Qy	354	LMDPIVTLFGSVHENPMPOAGGMRSTLFPNFGSMFSVTDQ--HAKNEQWD---EENLHRD	408
Db	295	LKDPLVNLFGSLHEKMPAAGNTRSGIFPHFGSMFSTTADAPHGKPAHWKEDIESHYNKD	354
Qy	409	DEEVASDGAGGDYED--NLHSPLLRSQATGAEGKDIVHHGHRGSALSMMRRTLLGEGGD	465
Db	355	NDDYATDGDAGDDDDSDNDLRSPLMRSQTTSMD-KDMI PHPTSGSTLSMRRHSTLMQ-GN	412
Qy	466	GVSSTDIGGGWQLAWKWEKEGENGRKEGGFKRVYLHQEGVPGSRRGSI VSLPGGDVFE	525
Db	413	GESSMGIGGWHMGYRYENDE-----YKRYYLKEDGAE-SRRGSIISIPGGPD--G	460
Qy	526	GSEFVHAALVSQSALFSKGLAEPRMSDAAMVHPSEVAAGSRWKDLFEPGVRRALLVGV	585
Db	461	GGSYIHASALVRSVLGPKS----VHGSA MVPPKIAASGPLWSALLEPGVKRALVGV	515
Qy	586	GIQILQEFACINGVLYYTPQILEQAGVAVILSKFGLSSASASILISSLTTLMLPCIGFA	645
Db	516	GIQILQFSGINGVLYYTPQILERAGVDILLSSLGSSISASFISGLTTLLMLPAIIVA	575
Qy	646	MLLMDLSGRFLLJGTIPILIASLVILVSNLIDGLTAHALSTSVSVIVYFCFVMGF	705
Db	576	MRLMDVSGERSLLLWTIPVLIVSLVLVVISELLHISKVVNAALSTGCVLVYFCFFVMGY	635
Qy	706	PIPNI LCAEI PPTRV RGLCIA CAFTFWIGDIIVTYSLPVMINAIGLAGVFSIYAVVCLI	765
Db	636	PIPNI LCEI PPTRV RGLCIA ICAMVFWIGDIIVTYSLPVLLSSIGLVGVFSIYAACVI	695
Qy	766	SFVFVFLKVETKGMPLLEVITEFFFAVGAKQA AKA 800	
Db	696	SWIFVYMKVPETKGMPLLEVITYDYAFAGA-QAQASA 729	
RESULT 10			
ID	Q9SD00	PRELIMINARY;	PRT; 729 AA.
AC	Q9SD00;		
DT	01-MAY-2000	(TrEMBLrel. 13, Created)	
DT	01-MAY-2000	(TrEMBLrel. 13, Last sequence update)	
DT	05-JUL-2004	(TrEMBLrel. 27, Last annotation update)	
DE	Sugar transporter-like protein (Monosaccharide sensing protein		
DE	3).		
GN	Name=F26O13.130; Synonyms=mssp3;		
OS	Arabidopsis thaliana (Mouse-ear cress).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;		
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsi s.		
OX	NCBI_TaxID=3702;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Delseny M., Berger C., Cooke R., Grellet F., Laudie M., Mewes H.W.,		
RA	Lemcke K., Mayer K.F.X., Quetier F., Salanoubat M.;		

Query Match 55.8%; Score 2275; DB 2; Length 734;
Best Local Similarity 61.2%; Pred. No. 2.4e-139;
Matches 462: Conservative 111; Mismatches 146; Indels 3

RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Stamme C., Tjaden J., Trentmann O., Emmerlich V., Neuhaus E.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the sugar transporter family.
DR EMBL; AL133452; CAB63013.1; -.
DR EMBL; AJ532571; CAD58693.1; -.
DR PIR; T45780; T45780.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0005351; F: sugar porter activity; IEA.
DR GO; GO:0005215; F: transporter activity; IEA.
DR GO; GO:0008643; P: carbohydrate transport; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR003663; Sugar_transp.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF00083; Sugar_tr; 1.
DR PRINTS; PR00171; SUGTRNSPORT.
DR PROSITE; PS00850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Sugar transport; Transmembrane; Transport.
SQ SEQUENCE 729 AA; 78828 MW; 241A9415DE43FCCC CRC64;

Query Match 55.1%; Score 2245; DB 2; Length 729;
Best Local Similarity 60.4%; Pred. No. 2.2e-137;
Matches 457; Conservative 110; Mismatches 144; Indels 46; Gaps 14;

QY 56 MSGAVLVAIVASIGNLQGDWNTAIAAALYIKKFFQIQNEPTVEGLIVMSLIGATIVT 115
Db 1 MRSVVLVALAAAGNMLQGDWNTAIAAGAVIYIKKFFHLEKEPKIEGLIVAMSLIGATLIT 60

QY 116 TFSGLSDSIGRRPMLILSSILYFFSGLIMLWSPNVVLLARFVDGFGIGLAVTLVPLY 175
Db 61 TFSGPVSDKVGRRSMLILSSVLYFLSSIVFWFSNVVLLFARLLDGGFGLAVTLVPIY 120

QY 176 ISEIAPSEIRGLLNTLPQFSGGGMFLSYCMVFCMSLSPSPDWIRIMGLVLAIPSLFFFG 235
Db 121 ISETAPSEIRGLLNTLPQFSGGGMFLSYCLVFCMSLQSPSWRLMLGLVLSIPSIAFVL 180

QY 236 TIFPLPESPRWLVSQGRMAEAKKVLQKRGKDDVSGELSLLEGLVGGDTISIEEYIIGP 295
Db 181 AAFPLPESPRWLVSQGRMDEARQVLQRLRGREDVSGELALLVEGLVGKDTISIEEYVIGP 240

QY 296 ATEAADDLVTGDKEQITLYGPEEQSWIARPSKGPIMGLSVLSASRHGSMVNQSVPLM 355
Db 241 DNEENEGGNELPRKDQIKLYGPDGQSQWMAKPVKGQ----SSLASRQGSMLPRGGSLM 296

QY 356 DPIVTLFGSVHENMP--QAGGSMRSTLFPNFGSMFSTVDQHAQNEQWDEENLHRDDEYA 413
Db 297 DPLVTLFGSIHENLPSNNVASSRSMFLPNMGSLGMMGR--QESQNDPE---RNED-- 349

QY 414 SDGAGGDYEDNLHSPILSRQATGAEGKDIVHGHGSGALSMMRQTL-LGEGGDGVSTDI 472
Db 350 ----SSDQDENLNSPLSPQTT--EPDD--YHQRTVGTMHRRQSSLEMANVGETATATSI 401

QY 473 GGGWQLAWKWEKEGNGRK-EGGFKRVYLHQB-----GVPGSRGSIIVSL-PGGGDV 523
Db 402 GGGWQLAWKYNDKVGADKRVNGGLQRMVTHEETANNNTNNIPFSRRGSLLSFHPEGDGH 461

QY 524 FEGSEFVHAAALVSQALFSKGLAEPRMSDAAVMHPSEVAAKGSRWKDLFEPGVRRALL 583
Db 462 DQVNGYVQAAALVSQASMPGKGGETAML-----PKEV-KDGPGRRELKEPGVKRALMV 514

QY 584 GVGIIQLQQFAGINGVLYTTPQILEQAGVAVILSKFGLSSASASILISLTLTLLMLPCIG 643
Db 515 GVGIIQLQQFAGINGVLYTTPQILEETGVSSLLTNLGISAESASLISALTTLTLLMLPCI- 573

QY 644 FAMLMDLSGRRFLLLTIPILIASLVILVSNLIDLGTLAHALLSTVSIVVYFCCFVMG 703
Db 574 -----LVSMRSLMLSTIPILISLVTLVIGSLVNLGGSINALISTASVTYVLSFCFVMG 626

QY 704 FGPIPNILCAEIPFTRVRGICIAICAFTHWIGDIIVTYSLPVMLNAIGLAGVFSIYAVVC 763
Db 627 FGAI PNILCSEIFPTSVRGICITICALTFWICDIIVTTLTPVMLKSGIAGVFGIYAIVC 686

QY 764 LISFVFLKVPETKGMPLVITEFFAVGAKQAAKA 800
Db 687 AVAVFVFLKVPETKGMPLVISEFFSVGAKQODAAA 723

RESULT 11
Q8GT51
ID Q8GT51 PRELIMINARY; PRT; 753 AA.
AC Q8GT51;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Sugar transporter.
GN Name=stp2;
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC Triticeae; Hordeum.
OX NCBI_TaxID=4513;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Developing caryopsis;
RX MEDLINE=22424051; PubMed=12535352;
RA Weschke W., Panitz R., Gubatz S., Wang Q., Radchuk R., Weber H.,
RA Wobus U.;
RT "The role of invertases and hexose transporters in controlling sugar
RT ratios in maternal and filial tissues of barley caryopses during early
RT development.";
RL Plant J. 33:395-411(2003).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the sugar transporter family.
DR EMBL; AJ534446; CAD58959.1; -.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0005351; F: sugar porter activity; IEA.
DR GO; GO:0005215; F: transporter activity; IEA.
DR GO; GO:0008643; P: carbohydrate transport; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR003663; Sugar_transp.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF00083; Sugar_tr; 1.
DR PRINTS; PR00171; SUGTRNSPORT.
DR PROSITE; PS00850; MFS; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Sugar transport; Transmembrane; Transport.
SQ SEQUENCE 753 AA; 81489 MW; A533B2B341922A9B CRC64;

Query Match 53.8%; Score 2191.5; DB 2; Length 753;
Best Local Similarity 66.2%; Pred. No. 6.8e-134;
Matches 440; Conservative 76; Mismatches 132; Indels 17; Gaps 9;

QY 137 LYFFSGLIMLWSPNVVLLARFVDGFGIGLAVTLVPLVYIIEIAPSEIRGLLNTLPQFSG 196
Db 65 LVFLASNVMLWAPNVVLLARLIDGFGIGLAVTLVPLVYIIEIAPSEIRGLLNTLPQFSG 124

QY 197 SGGMFLSYCMVFGMSLSPSPDWIRIMGLVLAIPSLFFFGTLTIFYPSPRWLVSKGRMAE 256
Db 125 SGGMFLSYCMVFTMSLMPQPDWRIMGLVLSIPSLMYFALTIVFLPSPRWLVSKGRMAE 184

QY 257 KKVLQKLRGKDDVSGELSLLEGLVGGDTISIEEYIIGPATEAADLVTGDKEQITLYG 316
Db 185 KRVLQRLRGREDVSGEMALLVEGLV-GKTHIEEYIIGPDDELADGLAP-DQEKLKLYG 242

QY 317 PEEGQSWIARPSK--GPIMLGSVLSASRHGSMVNQSVPLMDPIVTLFGSVHENMPQAG 374

Db 243 AEEGWSIARPVRRGGQSGALGSLGMSRHGSMVSGKSLVDPVLTFLGVSHERKPEVMG 302
QY 375 SMRSTLFPNFGSMFVTDQHAQNEQWDEENLHRDDEEYADGAGGDEYNLHSPILSRQA 434
Db 303 SMRSTLFPNFGSMFVAEQQAKADWDAES-HRDEDEDYADHGADDIEDNLSPLISRQA 361
QY 435 TGAEGKDIV--HHGHRGSALSRRQTLLEGGDGVSSDTIGGQWQALAKWSEKEGENGRK 492
Db 362 TSVEGKEIAAPHGSIIMGVSESSMQ-----GGDAVSSMGIGGQWQALAKWTEREGADGRK 416
QY 493 EGGFKRVYLHQEGVPGSRRGSIIVSLPGGGDVFESEFVHAAALVSQSALFSKGLABPRMS 552
Db 417 EGGAQRIYLHEEGVSGDRRGSIILSMP-GGDIPPGGEYIQAAALVSQALYSKDLTEQQLA 475
QY 553 DAAMVHPSEVAAKGSRWKDLFEPGVRRALLVGVGIQILQFAGINGVLYYTPQILLEQAGV 612
Db 476 GPAMVHPSEAVAKGKWAELFEPGVKHALFVGIGLQILQFAGINGVLYYTPQILLEQAGV 535
QY 613 AVILSKFGLSSASASILISLTLMLPCIGFAMLLMDLSGRFLLLTIPILIASLVIL 672
Db 536 GILLSNICLSSSSASILISALTLLMLGYISDRIGARAITAASFMYCSIPALFFFYRRAIL 595
QY 673 VVSNLIDLGLTAHALLSTVSVIVVFCFFVMGPGPIP--NILCABIFPTRVRGLCIAICAF 730
Db 596 VLVNVLDTGTMVHAALSTISVIVVFCFFVM--GPIPEGSTSSVKIDTDRANAASLAKALL 653
QY 731 TFWIGDIIVYTLPLVMLNAIGLAGVFCGIYAVVCMIAFVFMKVPETKGMPLVITEFFA 790
Db 654 TFWIGDIIVYTLPLVMLNAIGLAGVFCGIYAVVCMIAFVFMKVPETKGMPLVITEFFS 713
QY 791 VGAKQ 795
Db 714 VGAKQ 718

RESULT 12
Q96290
ID Q96290 PRELIMINARY; PRT; 734 AA.
AC Q96290;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Sugar transporter.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Tjaden J., Neuhaus E.;
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the sugar transporter family.
DR EMBL; Z50752; CAA90628.1; -.
DR PIR; T51139; T51139.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005351; F:sugar porter activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0008643; P:carbohydrate transport; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR003663; Sugar_transpt.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF00083; Sugar_tr; 1.
DR PRINTS; PR00171; SUGTRNSPORT.
DR PROSITE; PS00850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Sugar transporter; Transmembrane; Transport.
SQ SEQUENCE 734 AA; 79462 MW; 1B009F5F457AF6E1 CRC64;

Query Match 53.6%; Score 2183; DB 2; Length 734;
Best Local Similarity 59.5%; Pred. No. 2.4e-133;
Matches 449; Conservative 110; Mismatches 160; Indels 36; Gaps 14;
QY 56 MSGAVLVAIVASIGNLLQGDWNTAAAVLYIKKEFQLQNEPT-VEGLIVSMSLIGATIV 114
Db 1 MKGATLVALAATIGNFLOGWDNATAGAMVYINKDNL---PTSQGLVAVMSLIGATIVI 57
QY 115 TTFSGPLSDSICRRPMLILSSILYFFESGLIMLWSPNVYVLLLARFVDGFGIGLAVTLVPL 174
Db 58 TTCSGPISDWLGRRPMLILSSVMYFVCGLIMLWSPNVYVLCFARLNGFGAGLAVTLVPV 117
QY 175 YISEIAPSEIRGLLNTLPQFSGSGGMFELSYCMVFGMSLSPSPDWRIMLGVLAIPSLFFFG 234
Db 118 YISETAPPEIRGQLNTLPQFLGSGGMFELSYCMVFTMSLSDSPSWRAMLGVLISPLLYLF 177
QY 235 LTIFYLPESPRWLVSQGRMAEAKKVLQKLRGKDDVSGELSLLEGLEVGDDTSIEEYIIG 294
Db 178 LTVFYLPESPRWLVSQGRMDEAKRVLQQLCGREDVTDEALLVEGLDIGEKTMEDLLVT 237
QY 295 PATEAADLVTDCKE-QITLYGPEEGSQWIARPSKGPIMGLSVLSLASRHGSMVNQSV 353
Db 238 LEDHEGDDTLETVDDEGQIRLYGTHENQSYLARVPEQ---NSSLGLRSRHGSLANQSMI 294
QY 354 LMDPIVTLFGSVHENMPQAGGSMRSTLFPNFGSMFSTVDQ--HAKNEQWD---EENLHRD 408
Db 295 LKDPLVNLFGSLHEKMPGAGGNTRSGIFPHFGSMFSTTADAPHGKPAHWEKDIESHYNKD 354
QY 409 DEEYASDGAGDYED---NLHSPLLSRQATCAEGKDIVHHGHRGSALSRRQTLLEGGD 465
Db 355 NDDYATDDGAGDDSDNDRSPLMSRQTSMQ--KDMIPHPTSGTSLSMRRHSTLMQ--GN 412
QY 466 GVSSTDIGGQWQALAKWSEKEGENGRKEGFKRVYLHQEGVPGSRRGSIIVSLPGGGDVFE 525
Db 413 GESSMGIGGQWGMGYRYENDE-----YKRYLKDGAEE-SRRGSIISIPGGPD--G 460
QY 526 GSEFVHAAALVSQSALFSKGLAEPRMSDAAAMVHPSEVAAKGSRWKDLFEPGVRRALLVGV 585
Db 461 GGSYIHASALVSRVSLGPKS-----VHGSAMVPPPEKIAASGPLWSALLEPQVKRALVGV 515
QY 586 GIQILQOFAGINGVLYYTPQILEQAGVAVILSKFGLSSASASILISLTLMLPCIGFA 645
Db 516 GIQILQOFSGINGVLYYTPQILERAGVDILLSSGLSSISASFLISGLTLLMLPAIVVA 575
QY 646 MLLMDLSGRRFLLLTIPILIASLVILVSNLIDLGLTAHALLSTVSVIVVFCFFVMGFG 705
Db 576 MRLMDVSGRRSLLLTIPVLIVSLVVLVISELIHISKVNAALSTGCVVLYFCFFVMGYG 635
QY 706 PIPNILCAEIFFPTRVRGLCIAICAFTEFWIGDIIIVTSLPVMNLNAIGLAGVFSIYAVVCLI 765
Db 636 PFQTSVLKSSQQADRGLCIAICAMVFWIGDIIIVTSLPVLSSIELVGVFSIYAAVCVI 695
QY 766 SFVVFVFLKVPETKGMPLVITEFFFAVGAKQAAAKA 800
Db 696 SWIFVYMKVPETKGMPLVITDYFAFGA-QAQASA 729

RESULT 13
Q6K967
ID Q6K967 PRELIMINARY; PRT; 652 AA.
AC Q6K967;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative hexose transporter.
GN Name=OJ1149_C12.19;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.

RA Sasaki T., Matsumoto T., Yamamoto K.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the sugar transporter family.
DR EMBL; AP004082; BAD23011.1; -.
DR InterPro; IPR000566; Lipocln_cytFABP.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR003663; Sug_transporter.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF00083; Sugar_tr; 1.
DR PRINTS; PR00171; SUGTRNSPORT.
DR PROSITE; PS00213; LIPOCALIN; UNKNOWN_1.
DR PROSITE; PS50850; MFS; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Sugar transport; Transmembrane; Transport.
SQ SEQUENCE 652 AA; 68827 MW; EEE20446D2F9B1F6 CRC64;

Query Match 48.2%; Score 1963; DB 2; Length 652;
Best Local Similarity 55.1%; Pred. No. 4.2e-119;
Matches 408; Conservative 95; Mismatches 134; Indels 104; Gaps 12;

QY 56 MSGAVLVAIVASIGNLLQGDWNTAIAAAVLYIKKEF-QLQNEPTVEGLIVMSLIGATIV 114
DB 1 MRGAVVAAAAAAGVNMQLQGDWNTAIAAGALLYMRRDLPALQAHPALQGLVATSLIGATIV 60

QY 115 TTFSGPLSDSIGRRPMLILSSILYFFSGLIMLWSPNVVLLARFVDGFGIGLAVTLVPL 174
DB 61 TTFSGPLSDSIGRRPMLIASALLYSLAGLLMLWSPNPVILLARLVDGFGAIGLAVTLVPV 120

QY 175 YISEAPSEIRGLLNTLPQFSGGGMFLSYCMVFGMSLSPSPDWRIMLGVLAIPLSLFFFG 234
DB 121 YISETAPPDTRGLLNTLPQLTGTGMFLSYCMVFLITLAPIPNWRLMLGVLLPALLYLL 180

QY 235 LTIFYPESPRLVSKGRMAEAKKVLQKRGKDDVSGELSLLLLEGLEVGGDTSEIYYIG 294
DB 181 LTIFFLPESPRLVSKGRMKEARTVLQMLRGRQDVSAMALLVEGLTTGRDTAIEEYVVG 240

QY 295 PATEAADDLVTDGKEQITLYGPEEGQSWIARPSKGPIMLGSVLSLASHGSMVNQSVPL 354
DB 241 PTDEA-----KVTLYG---GMSSGLAPGS---MFGSAVTLASRQGSMLDH---L 280

QY 355 MDPIVTLFGSVHENPMQAGSMRSTLFPNFGSMFVTDQAKNEQWDEENLHRDDEEYAS 414
DB 281 KDPVVALDLSLHDVNPAGG---TTDVPNLGSMIGVHDPID--WDEEN----- 325

QY 415 DGAGGDYEDNLHSPLLSRQATGAEGKDIVHHGHRGSALSMMRQTLLEGGDGVSSDTIGG 474
DB 326 ---SGDDDDGDIAPLLTM-----EGEAATSTVGIGG 353

QY 475 GWQLAWKWEKEGNGRKEGGFKRVYLHQEGVPGSRRGSIIVSLPGGDFVEGSEFVHAAA 534
DB 354 GWQLAWKWEKVAADGTRQSTVKRMYLHEEQAE-----VHAAA 392

QY 535 LVQSALFSKGLAEPRMSDAAMVHPSEVAAGSRWKDLFEP-GVRRALLVGVGIQILQQF 593
DB 393 LVQSALCTKKEAEVEEG-----WREVLPGGVRRHALVCGVAIQILQQF 438

QY 594 AGINGVLYTPQILEQAGVAVILSKFGLSSASILISLTLMLPCIGFAMLLMDLSG 653
DB 439 SGISGVLLYTPQILEQAGVGLLSRLGLRDDSASILISGVTTLMLPSIGVAMRLMDVSG 498

QY 654 RRFLLLTGTPILIASLVILVNSLIDLTGLAHALLSTVSIVVYFCCFVMGFGPIPNILCA 713
DB 499 RRSLLLTWITPLLVASLAVLVAASVAPMAAAHAAVCTGSVVVYLCFVMGFGPIPNILCA 558

QY 714 EIPFTRVRGLCIAICAFIFWIGDIIIVTSLPVMVNAIGLAGVFSIYAVVCLISFVFVFLK 773
DB 559 EIPFTRVRGLCIAICSLAPFWLADIATVYTLPMVSLASGLAGLFAIYAAVCCVALVFVALR 618

QY 774 VPETKGMPLVITEFFAVGAK 794
DB 619 VPETKGLPLEVIIDFFNVGAK 639

RESULT 14
Q8H887
ID Q8H887 PRELIMINARY; PRT; 714 AA.
AC Q8H887;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative sugar transporter protein.
GN Name=OJ1528D07.4;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Wing R.A., Yu Y., Soderlund C., Kim H.-R., Rambo T., Saski C.,
RA Currie J., Collura K.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the sugar transporter family.
DR EMBL; AC099739; AAN17390.1; -.
DR Gramene; Q8H887; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005351; F:sugar porter activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0008643; P:carbohydrate transport; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR003663; Sugar_transpt.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF00083; Sugar_tr; 1.
DR PRINTS; PR00171; SUGTRNSPORT.
DR PROSITE; PS50850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN_1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Sugar transport; Transmembrane; Transport.
SQ SEQUENCE 714 AA; 77882 MW; E6DA98A692106885 CRC64;

Query Match 43.3%; Score 1763; DB 2; Length 714;
Best Local Similarity 47.5%; Pred. No. 4.8e-106;
Matches 377; Conservative 125; Mismatches 193; Indels 98; Gaps 14;

QY 4 GSWLAVQTPFTPLDRERLLPSVVLALPGPLPPASCSSQEPVTSDDILEDKMSGAVLVA 63
DB 9 GAW-----KNTRQQYSPSPISLSSVFSSESLSTTTAD----GSTANAVLAA 54

QY 64 IVASIGNLLQGDWNTAIAAAVLYIKKEFQLQNEPTVEGLIVMSLIGATIVTFSGLSD 123
DB 55 IVASIGNLLQGDWNTAIAAGAIMYIKNEFNQNDPMMEGLILAMSLIGATITALSGMTN 114

QY 124 SIGRRPMLILSSILYFFSGLIMLWSPNVVLLARFVDGFGIGLAVTLVLYISEIAPSE 183
DB 115 SIGKRPLLSVAAILYSISALIMFQASNEYMELLARLIYFGSGLVVTYAPLYISETAPN 174

QY 184 IRGLLNTLPQFSGGGMFLSYCMVFGMSLSPSPDWRIMLGVLAIPLSLFFGLTIFYLPES 243
DB 175 MRGLLNTLPQFNGSLGMLLSYIMVFLMSLTUNPNWRIMLGSLSIPSFVLLLCIFYLPES 234

QY 244 PRWLVSGRMAEAKKVLQKRGKDDVSGELSLLLLEGLEVGGDTSEIYYIGPATEAADL 303
DB 235 PVFLVSKGKIEEAKVMKRLRGTVNEVSSEIAFLIQGLTVQDQNYIEDYIGHNDEFDD- 293

QY 304 VTDGDKEQITLYGPEEGQSWIARPSKGPIMLGSVLSLASHGSMVNQSVPLMDPIVTLFG 363
DB 294 QSIISNTETTKLYGHEEGVTWFARPFKGNV-----VESDHSPIPN-----LLDPVTLFD 343

QY 364 SVHENMPQAGGSMRSTLFPNFGSMFVTDQAKNEQWDEENLHRDDEEYASDGAGDYED 423
DB 344 SIHGN-----LNTPEFTSSGNMSNDIEQ-PKTDLESQEDLDT-----DYED 384

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2004, 11:33:18 ; Search time 55.5861 Seconds
(without alignments)
3568.829 Million cell updates/sec

Title: US-10-051-909-36
Perfect score: 2779
Sequence: 1 PSSSSSFRPAGKKKKKNQG.....TIFVSLSIQRQLWLPECLS 553

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2777	99.9	553	8	ADG47940 Corn Beta
2	1801.5	64.8	501	8	ADG47942 Beta-vulg
3	951.5	34.2	437	3	AAG13750 Arabidops
4	951.5	34.2	447	3	AAG13749 Arabidops
5	951.5	34.2	463	3	AAG13748 Arabidops
6	943.5	34.0	437	3	AAG20978 Arabidops
7	943.5	34.0	447	3	AAG20977 Arabidops
8	943.5	34.0	463	3	AAG20976 Arabidops
9	933	33.6	438	3	AAG38871 Arabidops
10	933	33.6	448	3	AAG38870 Arabidops
11	933	33.6	464	3	AAG38869 Arabidops
12	891	32.1	481	8	ADM48307 Polypepti
13	874.5	31.5	468	3	AAG39282 Arabidops
14	874.5	31.5	470	3	AAG39281 Arabidops
15	868.5	31.3	474	3	AAG50066 Arabidops
16	853.5	30.7	472	3	AAG50064 Arabidops
17	853.5	30.7	474	3	AAG50063 Arabidops
18	829.5	29.8	415	3	AAG50067 Arabidops
19	824	29.7	477	3	AAG30433 Arabidops
20	824	29.7	483	3	AAG30432 Arabidops
21	819	29.5	420	3	AAG30434 Arabidops
22	807.5	29.1	415	3	AAG50065 Arabidops
23	805.5	29.0	471	3	AAG48646 Arabidops
24	805.5	29.0	558	3	AAG48645 Arabidops
25	805.5	29.0	564	3	AAG48644 Arabidops

26	801.5	28.8	398	3	AAG39283 Arabidops
27	790.5	28.4	454	3	AAG37675 Arabidops
28	790.5	28.4	459	3	AAG37674 Arabidops
29	786.5	28.3	454	3	AAG06925 Arabidops
30	786.5	28.3	460	3	AAG06924 Arabidops
31	774.5	27.9	409	3	AAG37676 Arabidops
32	770.5	27.7	409	3	AAG06926 Arabidops
33	758.5	27.3	382	3	AAG10665 Arabidops
34	741.5	26.7	375	3	AAG10666 Arabidops
35	738.5	26.6	384	3	AAG50068 Arabidops
36	717.5	25.8	369	3	AAG10667 Arabidops
37	688	24.8	489	4	ABB63511 Drosophil
38	682.5	24.6	370	3	AAG09696 Arabidops
39	681.5	24.5	368	3	AAG09697 Arabidops
40	571	24.1	477	4	AAB66934 Murine GL
41	668.5	24.1	478	4	AAB66933 Rat GLUTX
42	661.5	23.8	478	4	AAB66939 GLUTX1 co
43	658.5	23.7	477	4	AAE04888 Human tra
44	658.5	23.7	477	6	ABP58364 Human sol
45	652.5	23.5	477	4	AAB66932 Human GLU

ALIGNMENTS

RESULT 1
ADG47940
ID ADG47940 standard; protein; 553 AA.
XX
AC ADG47940;
XX
XX
DT 11-MAR-2004 (first entry)
XX
DE Corn Beta-vulgaris-like sugar transport protein #4.
XX
KW Arabidopsis-like sugar transport protein;
KW Beta-vulgaris-like sugar transport protein; transgenic; physical mapping;
KW corn; plant.
XX
OS Zea mays.
XX
FH Key Location/Qualifiers
FT Misc-difference 528
FT /note= "Encoded by AGN"
XX
PN US2002199217-A1.
XX
PD 26-DEC-2002.
XX
PF 17-JAN-2002; 2002US-00051909.
XX
PR 24-APR-1998; 98US-0083044P.
PR 14-APR-1999; 99US-00291922.
XX
PA (HELE/) HELENTJARIS T G.
XX
PI Helentjaris TG;
XX
DR WPI; 2004-040967/04.
DR N-PSDB; ADG47939.
XX
PT New isolated polynucleotide encoding a polypeptide having sugar transport
PT protein activity, for producing a transformed plant and for use as probes
PT in physical mapping.
XX
PS Claim 18; SEQ ID NO 36; 71pp; English.
XX
CC The invention relates to Arabidopsis-like or Beta-vulgaris-like sugar
CC transport proteins and their corresponding nucleic acid sequences. The
CC sequences of the invention are useful to transform a cell. These are also
CC useful to produce a transgenic plant. Probes derived from sequences
CC encoding sugar transport protein may be used for physical mapping. The
CC present sequence is corn Beta-vulgaris-like sugar transport protein.

XX Sequence 553 AA;
SQ
Query Match 99.9%; Score 2777; DB 8; Length 553;
Best Local Similarity 100.0%; Pred. No. 6.6e-250;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PSSSSFRPAGKKKKKQGLRREAVPGRPASELRTRVMGGSGNRGGAGAGEESGSDHDG 60
Db 1 PSSSSFRPAGKKKKKQGLRREAVPGRPASELRTRVMGGSGNRGGAGAGEESGSDHDG 60
QY 61 VLRRPLNTGWSYRMSSRQSSFAFGSLSNVGMVGAIAFGMAEYIGRKGLMIAAIPNIIGW 120
Db 61 VLRRPLNTGWSYRMSSRQSSFAFGSLSNVGMVGAIAFGMAEYIGRKGLMIAAIPNIIGW 120
QY 121 SPTQDAMVRDNLNLSISEFSAFGSLSNVGMVGAIAFGMAEYIGRKGLMIAAIPNIIGW 180
Db 121 SPTQDAMVRDNLNLSISEFSAFGSLSNVGMVGAIAFGMAEYIGRKGLMIAAIPNIIGW 180
QY 181 LAISFAKDASFLYMGRLLEGFGVGIISYTPVYIAEISPNMRGALGSVNQLSVTFGIFL 240
Db 181 LAISFAKDASFLYMGRLLEGFGVGIISYTPVYIAEISPNMRGALGSVNQLSVTFGIFL 240
QY 241 AYLLGMFIPWRLLAIVIGALPCTMLIPGLFFIPESPRWLAKMNLTEDCETSLQVLRGFETD 300
Db 241 AYLLGMFIPWRLLAIVIGALPCTMLIPGLFFIPESPRWLAKMNLTEDCETSLQVLRGFETD 300
QY 301 ITTEVNDIKRAVASSSKRTTISFQELNQKKYRTPLLLGIGLLVLQNLGSLGVVFFL 360
Db 301 ITTEVNDIKRAVASSSKRTTISFQELNQKKYRTPLLLGIGLLVLQNLGSLGVVFFL 360
QY 361 FXAAGVNTSDLATCSLGAIQVLTATGVTWLLDRAGRRIILLIISTSGMTLCLLAVSVVFFL 420
Db 361 FXAAGVNTSDLATCSLGAIQVLTATGVTWLLDRAGRRIILLIISTSGMTLCLLAVSVVFFL 420
QY 421 KDNISQDSNSYIILTMISLVGIVSFVITFSFGMGAIPLMMSEILPVSIKSLGSIATLA 480
Db 421 KDNISQDSNSYIILTMISLVGIVSFVITFSFGMGAIPLMMSEILPVSIKSLGSIATLA 480
QY 481 NWLTSFAITMTNMLTWSVGGTFLSYMVVSAFTIVFVVLWVPETKXNSRGDTIFVSL 540
Db 481 NWLTSFAITMTNMLTWSVGGTFLSYMVVSAFTIVFVVLWVPETKXNSRGDTIFVSL 540
QY 541 IQRQLQWLPECLS 553
Db 541 IQRQLQWLPECLS 553

RESULT 2
ADG47942
ID ADG47942 standard; protein; 501 AA.
XX AC ADG47942;
XX DT 11-MAR-2004 (first entry)
XX DE Beta-vulgaris-like sugar transport protein #2.
XX KW Arabidopsis-like sugar transport protein;
XX KW Beta-vulgaris-like sugar transport protein; transgenic; physical mapping;
XX KW plant; beetroot.
XX OS Beta vulgaris.
XX PN US2002199217-A1.
XX PD 26-DEC-2002.
XX PF 17-JAN-2002; 2002US-00051909.
XX PR 24-APR-1998; 98US-0083044P.
XX PR 14-APR-1999; 99US-00291922.
XX

PA (HELE/) HELENTJARIS T G.
XX Helentjaris TG;
XX WPI; 2004-040967/04.
PT New isolated polynucleotide encoding a polypeptide having sugar transport
PT protein activity, for producing a transformed plant and for use as probes
PT in physical mapping.
XX Example 4; SEQ ID NO 38; 71pp; English.
XX The invention relates to Arabidopsis-like or Beta-vulgaris-like sugar
CC transport proteins and their corresponding nucleic acid sequences. The
CC sequences of the invention are useful to transform a cell. These are also
CC useful to produce a transgenic plant. Probes derived from sequences
CC encoding sugar transport protein may be used for physical mapping. The
CC present sequence is wheat Beta-vulgaris-like sugar transport protein.
XX Sequence 501 AA;
Query Match 64.8%; Score 1801.5; DB 8; Length 501;
Best Local Similarity 72.2%; Pred. No. 6.2e-159;
Matches 351; Conservative 62; Mismatches 64; Indels 9; Gaps 2;
QY 51 GEESGSDHDG-----VLRRLNTGWSYRMSSRQSSAPGT---SSMAVLRSHVSAFL 101
Db 5 GEESGSDGGRRTASADLRKPFHTGWSYRMSSAGGGGMSRLGSSAYSLRDSVSAVL 64
QY 102 CTLIVALGPIQFGTSGFSSPTQDAMVRDNLNLSISEFSAFGSLSNVGMVGAIAFGMAE 161
Db 65 CTLIVALGPIQFGTSGFSSPTQDAIISDLGLTSEFLGSLSNVGMVGAIAFGMAE 124
QY 162 YIGRKSLMIAAIPNIIGWLAISPAKDAFLYMGRLLEGFGVGIISYTPVYIAEISPN 221
Db 125 YIGRKSLMIAAIPNIIGWLAISPAKDAFLYMGRLLEGFGVGIISYTPVYIAEISPN 184
QY 222 MRGALGSVNQLSVTFGIFLAIVIGALPCTMLIPGLFFIPESPRWLAKM 281
Db 185 MRGALGSVNQLSVTFGIFLAIVIGALPCTMLIPGLFFIPESPRWLAKM 244
QY 282 NLTEDCETSLQVLRGFETDITTEVNDIKRAVASSSKRTTISFQELNQKKYRTPLLLGIGL 341
Db 245 GKMEDFESSLQVLRGFETDIAVEVNEIKRSVQSSRRRTTIRFADIKQKRYSVPLMVGIGL 304
QY 342 LVLQNLGSLGVVFFLAVASSIFKAAGVTNSDLATCSLGAIQVLTATGVTWLLDRAGRRIILLI 401
Db 305 LVLQNLGSLGVVFFLAVASSIFKAAGVTNSDLATCSLGAIQVLTATGVTWLLDRAGRRIILLI 364
QY 402 ISTSGMTLCLLAVSVVFFLKDNLISQDSNSYIILTMISLVGIVSFVITFSFGMGAIPLMM 461
Db 365 ISTSGMTITLVVSVVFFLKDNLISQDSNSYIILTMISLVGIVSFVITFSFGMGAIPLMM 424
QY 462 SEILPVSIKSLGSIATLANWLTSFAITMTNMLTWSVGGTFLSYMVVSAFTIVFVVLW 521
Db 425 SEILPVNIKSLAGSVATLANWLTAWLITMTASMLSWNSGGTFAIYAAVCAGTLVFCVCLW 484
QY 522 VPETKG 527
Db 485 VPETKG 490

RESULT 3
AAG13750
ID AAG13750 standard; protein; 437 AA.
XX AC AAG13750;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 13359.
XX KW Protein identification; signal transduction pathway; metabolic pathway;

KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
OS EP1033405-A2.
PN 06-SEP-2000.
XX 25-FEB-2000; 2000EP-00301439.
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.

PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.

PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 34.2%; Score 951.5; DB 3; Length 437;
Best Local Similarity 42.3%; Pred. No. 1.1e-79;
Matches 181; Conservative 102; Mismatches 140; Indels 5; Gaps 2;
QY 100 FLCTLI VALGPIQFTSGFSPTQDAMVRDLNLSISEFSAFGLSNVGMVGAIASGQM 159
Db 3 YLSTFVAVCGPFAFGSCAGYSSPAQAAIRNDLSLTIAEFSLFGLTFTGAMIGAITSGPI 62
QY 160 ABYIGRKGSMLIAIPNIIGWLAIISFAKDAFLYMGRLLEGFGVGIIISYTPVPIAEISP 219
Db 63 ADLVGRKGAMRVSSAFVVGWLAIIFAKGVVALDLGLRLATGYGMGAFSVVPIFAEIAIP 122
QY 220 QNMRGALGSVNQLSVTFGIFLAYLLGMFIPWRLAVIGALPCTMLIPGLFFIPESPRWLA 279
Db 123 KTFRGALTTLNQILICTGVSVSFIIGTLVTRVLALIGIIPCAASFLGLFFIPESPRWLA 182
QY 280 KMNLTEDCETSLQVLRGFETDITTEVNDIKRAVASSSKRTTISFOELNQKKYRTPLLLGI 339
Db 183 KVGRDTEFEALRLRGKKADISEAAEIQDYIETLERLPKAKMLDLFQRRYIRSVLIAF 242
QY 340 GLLVLQNLSGINGVLFYASSIFKAAGVTNSDLATCSLGAIQVLATGVTTWLLDRAGRRIL 399
Db 243 GLMVFFQFGGINGICFYTSSIFEQAGFP-TRLGMIYAVLQVITALNAPIVDRAGRKPL 301
QY 400 LIISTSGMTLCLLAVSVVFFLKNISQDSNSYYILTMISLVGIVSFVITFSFGMGAIPWL 459
Db 302 LLVSATGLVIGCLIAAVSFYLVKH----DMAHEAVPVLAVVGIMVYIGSFSGMGAMPV 357

QY 460 MMSEILPVSTKSLGGSIAITLANWLTSPAITMTTNMLTWSVGGTFLSYMVVSAFTIVFV 519
Db 358 VMSEIFPINKGVAGGMATLVNWFEGAWAVSYTFNFLMSWSSYGTFLIYAALAIIVFVI 417
QY 520 LMVPETKG 527
Db 418 AIVPETKG 425
RESULT 4
AAG13749
ID AAG13749 standard; protein; 447 AA.
XX AC AAG13749;
XX 17-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 13358.
XX protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
OS Arabidopsis thaliana.
XX EP1033405-A2.
PN 06-SEP-2000.
PD 25-FEB-2000; 2000EP-00301439.
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.

PR	10-AUG-1999;	99US-0148171P
PR	11-AUG-1999;	99US-0148319P
PR	12-AUG-1999;	99US-0148341P
PR	13-AUG-1999;	99US-0148565P
PR	13-AUG-1999;	99US-0148684P
PR	16-AUG-1999;	99US-0149368P
PR	17-AUG-1999;	99US-0149175P
PR	18-AUG-1999;	99US-0149426P
PR	20-AUG-1999;	99US-0149722P
PR	20-AUG-1999;	99US-0149723P
PR	20-AUG-1999;	99US-0149929P
PR	23-AUG-1999;	99US-0149902P
PR	23-AUG-1999;	99US-0149930P
PR	25-AUG-1999;	99US-0150566P
PR	26-AUG-1999;	99US-0150884P
PR	27-AUG-1999;	99US-0151065P
PR	27-AUG-1999;	99US-0151066P
PR	27-AUG-1999;	99US-0151080P
PR	30-AUG-1999;	99US-0151303P
PR	31-AUG-1999;	99US-0151438P
PR	01-SEP-1999;	99US-0151930P
PR	07-SEP-1999;	99US-0152363P
PR	10-SEP-1999;	99US-0153070P
PR	13-SEP-1999;	99US-0153758P
PR	15-SEP-1999;	99US-0154018P
PR	16-SEP-1999;	99US-0154039P
PR	20-SEP-1999;	99US-0154779P
PR	22-SEP-1999;	99US-0155139P
PR	23-SEP-1999;	99US-0155486P
PR	24-SEP-1999;	99US-0155659P
PR	28-SEP-1999;	99US-0156458P
PR	29-SEP-1999;	99US-0156596P
PR	04-OCT-1999;	99US-0157117P
PR	05-OCT-1999;	99US-0157753P
PR	06-OCT-1999;	99US-0157865P
PR	07-OCT-1999;	99US-0158029P
PR	08-OCT-1999;	99US-0158232P
PR	12-OCT-1999;	99US-0158369P
PR	13-OCT-1999;	99US-0159293P
PR	13-OCT-1999;	99US-0159294P
PR	13-OCT-1999;	99US-0159295P
PR	14-OCT-1999;	99US-0159329P
PR	14-OCT-1999;	99US-0159330P
PR	14-OCT-1999;	99US-0159331P
PR	14-OCT-1999;	99US-0159637P
PR	14-OCT-1999;	99US-0159638P
PR	18-OCT-1999;	99US-0159584P
PR	21-OCT-1999;	99US-0160741P
PR	21-OCT-1999;	99US-0160767P
PR	21-OCT-1999;	99US-0160768P
PR	21-OCT-1999;	99US-0160770P
PR	21-OCT-1999;	99US-0160814P
PR	21-OCT-1999;	99US-0160815P
PR	22-OCT-1999;	99US-0160980P
PR	22-OCT-1999;	99US-0160981P
PR	22-OCT-1999;	99US-0160989P
PR	25-OCT-1999;	99US-0161404P
PR	25-OCT-1999;	99US-0161405P
PR	25-OCT-1999;	99US-0161406P
PR	26-OCT-1999;	99US-0161359P
PR	26-OCT-1999;	99US-0161360P
PR	26-OCT-1999;	99US-0161361P
PR	28-OCT-1999;	99US-0161920P
PR	28-OCT-1999;	99US-0161992P
PR	28-OCT-1999;	99US-0161993P
PR	29-OCT-1999;	99US-0162142P

```
Query Match      34.2%; Score 951.5; DB 3; Length 447;
Best Local Similarity 42.3%; Pred. No. 1.1e-79;
Matches 181; Conservative 102; Mismatches 140; Indels 5; Gaps 2;
```

Qy 100 FLCTLRVALGPIQFGFTSGFSSPTQDAMVRDLNLSISEFSAFGLSLNVGMVGAIASGQM 159

Db 13 YLSTFVAVCGPPAFGSCAGYSSPAQAARNDLSLTIAEFSLFGSLLTFGAMIGAITSGPI 72
QY 160 AEYIGRKGSMLTAAIPNIIGWLAIISPAKDASFLYMGRLLEGFGVGIISYTPVYIAEISP 219
Db 73 ADLVGRKGAMRVSSAFVCVVGWLAIIFAKGVVADLDLGRLATGYGNGAFSYVPIFIAEIAF 132
QY 220 QNMRGALGSVNQLSVTFEGIFLAYLLGMFIPWRLLAVIGALPCTMLIPGLFFIPESPRWLA 279
Db 133 KTFRGALTTLNQILICTGVSVSFIIGTLVTRVRLALIGIIPCAASFLGLFFIPESPRWLA 192
QY 280 KMNLTEDCETSLQVLRGFPETDITTEVNDIKRAVASSSKRTTISFQELNQKKYRTPLLLGI 339
Db 193 KVGRTDTEFEAALRKLRGKKADI SEEAABIQDIETLERLPKAKMLDLFORRYIRSVLIAF 252
QY 340 GLLVLQNLSGINGVLFYASSIFKAAGVNTNSDLATCSLGAIQVLATGVTTWLLDRAGRIL 399
Db 253 GLMVFAQFGGICFCYTSSIFEQAGFP-TRLGMIYAVLQVVITALNAPIVDRAGRKPL 311
QY 400 LIISTSGMTLCLLAVSVVFFLKDNIQSNSYYILTMISLVGIVSFVITFSFGMGAIPWL 459
Db 312 LLVSATGLVIGCLIAAASVFLKVH---DMAHEAVPVLAVGIMVYIGSFSAGMGAMPVW 367
QY 460 MMSEILPVSISKLSGGSIATLANWLTSFAITMTTNLMLTWSVCGTFLSYMVVSFAFTVFWV 519
Db 368 VMSEIFFPINIKGVAGGMATLVNWFAGAWAVSYTFNFLMSNSSYGTFLIYAAINALAIVFVI 427
QY 520 LWVPETKG 527
Db 428 AIVPETKG 435

RESULT 5
AAG13748
ID AAG13748 standard; protein; 463 AA.
XX
AC AAG13748;
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 13357.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.

PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.

PR	22-JUL-1999;	99US-0145192P;
PR	23-JUL-1999;	99US-0145145P;
PR	23-JUL-1999;	99US-0145218P;
PR	23-JUL-1999;	99US-0145224P;
PR	26-JUL-1999;	99US-01452276P;
PR	27-JUL-1999;	99US-0145913P;
PR	27-JUL-1999;	99US-0145918P;
PR	27-JUL-1999;	99US-0145919P;
PR	28-JUL-1999;	99US-0145951P;
PR	02-AUG-1999;	99US-0146386P;
PR	02-AUG-1999;	99US-0146388P;
PR	02-AUG-1999;	99US-0146389P;
PR	03-AUG-1999;	99US-0147038P;
PR	04-AUG-1999;	99US-0147204P;
PR	04-AUG-1999;	99US-0147302P;
PR	05-AUG-1999;	99US-0147192P;
PR	05-AUG-1999;	99US-0147260P;
PR	06-AUG-1999;	99US-0147303P;
PR	06-AUG-1999;	99US-0147416P;
PR	09-AUG-1999;	99US-0147493P;
PR	09-AUG-1999;	99US-0147935P;
PR	10-AUG-1999;	99US-0148171P;
PR	11-AUG-1999;	99US-0148319P;
PR	12-AUG-1999;	99US-0148341P;
PR	13-AUG-1999;	99US-0148565P;
PR	13-AUG-1999;	99US-0148684P;
PR	16-AUG-1999;	99US-0149368P;
PR	17-AUG-1999;	99US-0149175P;
PR	18-AUG-1999;	99US-0149426P;
PR	20-AUG-1999;	99US-0149722P;
PR	20-AUG-1999;	99US-0149723P;
PR	20-AUG-1999;	99US-0149929P;
PR	23-AUG-1999;	99US-0149902P;
PR	23-AUG-1999;	99US-0149930P;
PR	25-AUG-1999;	99US-0150566P;
PR	26-AUG-1999;	99US-0150884P;
PR	27-AUG-1999;	99US-0151065P;
PR	27-AUG-1999;	99US-0151066P;
PR	27-AUG-1999;	99US-0151080P;
PR	30-AUG-1999;	99US-0151303P;
PR	31-AUG-1999;	99US-0151438P;
PR	01-SEP-1999;	99US-0151930P;
PR	07-SEP-1999;	99US-0152363P;
PR	10-SEP-1999;	99US-0153070P;
PR	13-SEP-1999;	99US-0153758P;
PR	15-SEP-1999;	99US-0154018P;
PR	16-SEP-1999;	99US-0154039P;
PR	20-SEP-1999;	99US-0154779P;
PR	22-SEP-1999;	99US-0155139P;
PR	23-SEP-1999;	99US-0155486P;
PR	24-SEP-1999;	99US-0155659P;
PR	28-SEP-1999;	99US-0156458P;
PR	29-SEP-1999;	99US-0156596P;
PR	04-OCT-1999;	99US-0157117P;
PR	05-OCT-1999;	99US-0157753P;
PR	06-OCT-1999;	99US-0157865P;
PR	07-OCT-1999;	99US-0158029P;
PR	08-OCT-1999;	99US-0158232P;
PR	12-OCT-1999;	99US-0158369P;
PR	13-OCT-1999;	99US-0159293P;
PR	13-OCT-1999;	99US-0159294P;
PR	13-OCT-1999;	99US-0159295P;
PR	14-OCT-1999;	99US-0159329P;
PR	14-OCT-1999;	99US-0159330P;
PR	14-OCT-1999;	99US-0159331P;
PR	14-OCT-1999;	99US-0159637P;
PR	14-OCT-1999;	99US-0159638P;
PR	18-OCT-1999;	99US-0159584P;
PR	21-OCT-1999;	99US-0160741P;
PR	21-OCT-1999;	99US-0160767P;
PR	21-OCT-1999;	99US-0160768P;
PR	21-OCT-1999;	99US-0160770P;
PR	21-OCT-1999;	99US-0160814P;

[illegible]

RESULT 6
AAG20978
ID AAG20978 standard; protein; 437 AA.
XX
XX AAG20978;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 23367.
XX
KW protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.

XX PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 24-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 30-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.

PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 34.0%; Score 943.5; DB 3; Length 437;
Best Local Similarity 42.1%; Pred.No. 5.9e-79;
Matches 180; Conservative 102; Mismatches 141; Indels 5; Gaps 2;

QY 100 FLCTLIVLGIQFGFTSGFSSPTQDAMVRDLNLSISEFSAFGSLSNVGGMVGAIASGQM 159
Db 3 YLSTFVAVCGSFAFGSCAGYSPPAQAAIRNDLSLTIAEFSLFTFGAMIGAITSGPI 62
QY 160 AEYIGRKGSIMTAAIPNIIGWLAIISFAKDASFLYMGRLLEGFGVGIISYTPVYIAEISP 219
Db 63 ADLVGRKGAMRVSSAFVVGWLAIIFAKGVVALDLGLRLATGYGNGAFSVVPFIPIAEIAP 122
QY 220 QNMRGALGSVNQLSVTFGIFLAYLLGMETPWRLLAIVIGALPCTNMLIPGLFFIPESPRWLA 279
Db 123 KTFRGALTTLNQILICTGVSVSFIIGTLVTRVLAIGIIPCAASFLGLFFIPESPRWLA 182
QY 280 KMNLTEDCETSLQVLRGFETDITTEVNDIKRAVASSSKRTTISFOELNQKKYRTPLLLGI 339
Db 183 KVGRTDTEFEAALRKLGRKKADISEEAAEQDIETLERLPKAKMLDLFQRRYIRSVLIAF 242
QY 340 GLLVLQNLGSLGVLFYASSIFKAAGVTNSDLATCSLGAIQVLATGVTTLWLLDRAGRRIL 399
Db 243 GLMVFQQFGGINGICFYTSSEIFEQAGFP-TRIGMIIYAVLQVVITALNAPIVDRAGRKPL 301
QY 400 LIISTSGMTLCLLAVSVVVEFLKDNISQDSNSYYILTMISLVGIVSFVITFSFGMGAIPWL 459
Db 302 LILVSATGLVIGCLIAAVSFYLVKH---DMAHEAVPVLAVVGIMVYIGSFSGMGAMPWV 357
QY 460 MMSEILPVSIKSLGGSIIATLANWLTSFAITMTNMLTWSVGGTFLSYMVVSFAITVFEV 519
Db 358 VMSEIFPINIKGVAGGMATLVNWFAGAWAVSYTFNFLMSWSSYGTFLIYAINALAIVFEI 417
QY 520 LWPVETKG 527
Db 418 AIVPETKG 425

RESULT 7
AAG20977
ID AAG20977 standard; protein; 447 AA.
XX
AC AAG20977;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 23366.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.

PR	02-AUG-1999;	99US-0146388P;
PR	02-AUG-1999;	99US-0146389P;
PR	03-AUG-1999;	99US-0147038P;
PR	04-AUG-1999;	99US-0147204P;
PR	04-AUG-1999;	99US-0147302P;
PR	05-AUG-1999;	99US-0147192P;
PR	05-AUG-1999;	99US-0147260P;
PR	06-AUG-1999;	99US-0147303P;
PR	06-AUG-1999;	99US-0147416P;
PR	09-AUG-1999;	99US-0147493P;
PR	09-AUG-1999;	99US-0147935P;
PR	10-AUG-1999;	99US-0148171P;
PR	11-AUG-1999;	99US-0148319P;
PR	12-AUG-1999;	99US-0148341P;
PR	13-AUG-1999;	99US-0148565P;
PR	13-AUG-1999;	99US-0148684P;
PR	16-AUG-1999;	99US-0149368P;
PR	17-AUG-1999;	99US-0149175P;
PR	18-AUG-1999;	99US-0149426P;
PR	20-AUG-1999;	99US-0149722P;
PR	20-AUG-1999;	99US-0149723P;
PR	20-AUG-1999;	99US-0149929P;
PR	23-AUG-1999;	99US-0149902P;
PR	23-AUG-1999;	99US-0149930P;
PR	25-AUG-1999;	99US-0150566P;
PR	26-AUG-1999;	99US-0150884P;
PR	27-AUG-1999;	99US-0151065P;
PR	27-AUG-1999;	99US-0151066P;
PR	27-AUG-1999;	99US-0151080P;
PR	30-AUG-1999;	99US-0151303P;
PR	31-AUG-1999;	99US-0151438P;
PR	01-SEP-1999;	99US-0151930P;
PR	07-SEP-1999;	99US-0152363P;
PR	10-SEP-1999;	99US-0153070P;
PR	13-SEP-1999;	99US-0153758P;
PR	15-SEP-1999;	99US-0154018P;
PR	16-SEP-1999;	99US-0154039P;
PR	20-SEP-1999;	99US-0154779P;
PR	22-SEP-1999;	99US-0155139P;
PR	23-SEP-1999;	99US-0155486P;
PR	24-SEP-1999;	99US-0155659P;
PR	28-SEP-1999;	99US-0156458P;
PR	29-SEP-1999;	99US-0156596P;
PR	04-OCT-1999;	99US-0157117P;
PR	05-OCT-1999;	99US-0157753P;
PR	06-OCT-1999;	99US-0157865P;
PR	07-OCT-1999;	99US-0158029P;
PR	08-OCT-1999;	99US-0158232P;
PR	12-OCT-1999;	99US-0158369P;
PR	13-OCT-1999;	99US-0159293P;
PR	13-OCT-1999;	99US-0159294P;
PR	13-OCT-1999;	99US-0159295P;
PR	14-OCT-1999;	99US-0159329P;
PR	14-OCT-1999;	99US-0159330P;
PR	14-OCT-1999;	99US-0159331P;
PR	14-OCT-1999;	99US-0159637P;
PR	14-OCT-1999;	99US-0159638P;
PR	18-OCT-1999;	99US-0159584P;
PR	21-OCT-1999;	99US-0160741P;
PR	21-OCT-1999;	99US-0160767P;
PR	21-OCT-1999;	99US-0160768P;
PR	21-OCT-1999;	99US-0160770P;
PR	21-OCT-1999;	99US-0160814P;
PR	21-OCT-1999;	99US-0160815P;
PR	22-OCT-1999;	99US-0160980P;
PR	22-OCT-1999;	99US-0160981P;
PR	22-OCT-1999;	99US-0160989P;
PR	25-OCT-1999;	99US-0161404P;
PR	25-OCT-1999;	99US-0161405P;
PR	25-OCT-1999;	99US-0161406P;
PR	26-OCT-1999;	99US-0161359P;
PR	26-OCT-1999;	99US-0161360P;
PR	26-OCT-1999;	99US-0161361P;

	Query Match	34.0%;	Score 943.5;	DB 3;	Length 463;	
	Best Local Similarity	42.1%;	Pred. No. 6.4e-79;			
	Matches 180;	Conservative 102;	Mismatches 141;	Indels 5;	Gaps 2	
PR	28-OCT-1999;	99US-0161920P.				
PR	28-OCT-1999;	99US-0161992P.				
PR	28-OCT-1999;	99US-0161993P.				
PR	29-OCT-1999;	99US-0162142P.				
Qy	100	FLCTLIVALGP IQFGTSGFSSPTQDAMVRDLNLSISEFSAFGSLSNVGMVGAIASGQM	159			
Dd	29	YLSTFVAVCGSFAFGSCAGYSSPAQAARNDLSLTIAEFSLFGSLTFGAMIGAITSGPI	88			
Qy	160	AEYIGRKGSLMIAAIPNIIGWLAIISPAKDA SFLYMGRLLEGVFGIISYTVVPVYIAEISP	219			
Dd	89	ADLVGRKGAMRVSSAFVCVVGWLAIFAKGVVALDLGLRLATGYCGMAFSYVVPFI FIAEIAP	148			
Qy	220	QNMRGALGSVNQLSVTFGIFLAYLLGMFIPWRLLAVIGALPCTMLIPGLFFIPESPRWLA	279			
Dd	149	KTRFGALTINQILICTGVSVSFI GLTVTRVLALIGIIPCAASFGLGFIPESPRWLA	208			
Qy	280	KMNLTEDCETSLQVLRGFTDTITTEVNDIKRAVASSSKRTTISFQELNQKKYRTPLLLGI	339			
Dd	209	KVGRDTEFEAALRLRGKKADI SEEA EI QDYIETLERLPKAKMLDLFORRYIRSVLIAF	268			
Qy	340	GLJVLONLSGINGVLFYASSIFKAAGVTNSDLATCSLGAIQVLATGVTTWLLDRAGRRI L	399			
Dd	269	GLMVFOQFGGINGICFTSSIFEQAGFP -TRLGMI IYAVLQVVITALNAPIVD RAGRKPL	327			
Qy	400	LII STGMTCLLLAVSVVFELKDNISQDSNSYYILMTSLVCIVSFVITPSFGMGAIPWL	459			
Dd	328	LLVSATGLVIGCLIAAVSYFLKVH----DMAHEAVPV LAVGIMVYIGSF SAGMGAMPVW	383			
Qy	460	MMSEILPVSIKSLGGSATLANWLTSFAITMTTNLM L TW SVGGTFLSYM VVS AFTIVFV V	519			
Dd	384	VMSEIFPINIKGVAGGMATLVNMF GAWAVSYTENFNLMSWS SYGTFLI YAAINALAIVEVI	443			
Qy	520	LWVPETKG 527				
Dd	444	AIVPETKG 451				

PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144333P.

PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.

PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 33.6%; Score 933; DB 3; Length 438;
Best Local Similarity 42.0%; Pred. No. 5.7e-78;
Matches 180; Conservative 102; Mismatches 141; Indels 6; Gaps 3;

QY 100 FLCTLIVALGPIQFGFTSGSSPTQDAMVRDNLISSEFSAFGSLSNVGMVGAIASQM 159
DB 3 YLSTFVAVCGSFAGFCAGSSPAQAARNDLSLTIAEFLSFLGSLLTFFGAMIGAITSGPI 62
QY 160 AEVIGRKGSMLIAAIPNIIGWLAIISPAKADASFVYNGRLLEGFGVGIISYVPVYIAEISP 219
DB 63 ADLVGRKGAMRVSSAFVVGWLAIFAKGVVLDLGRLATGYGMGAFSYVVPVIFIAEIAIP 122
QY 220 QNMRGALGSVN-QLSVTFGIFLAYLLGMFIPWRLLAIVIGALPCTMLIPGLFFIPESPRWL 278
DB 123 KTFRGALTTLNQOILICTGVSVSFIIGTLVTRVRLALIGIIPCAASFLGLFFIPESPRWL 182
QY 279 AKMNLTEDCETSLQVLRGFETDITTEVNDIKRAVASSSKRTTISFQELNQKKYRTPLLG 338
DB 183 AKVGRDTEFEALRLKRGKADISEAAEIQDIETLERLPKAKMLDLFQRRYIRSVLIA 242
QY 339 IGLVLQNLGNGVLFYASSIFKAAAGVTNSDLATCSLGAIQVATGVTWLLDRAGRI 398
DB 243 FGLMVFFQFGINGICFTSSIFEQAGFP-TRLGMIYAVLQVVITALNAPIVDRAGRKP 301
QY 399 LLIISTSGMTCLLAVSVVFFLKDNIQSNSYVILTMISLVGIVSFVITFSFGMGAIPW 458
DB 302 LLLVSATGLVIGCLIAAVSYFLKVH---DMAHEAVPVLAUVGIMVYIGSFAGMGAMPW 357
QY 459 LMMSEILPVSIKSLGSLATLANWLTSFAITMTNMLTWSVGGTFLSYMVVSAFTIVFV 518
DB 358 VVMSEIFPINIKVAGGMATLVNWFAGAVSYTFNFLMSWSSYGTFLIYAANALAIIVFV 417
QY 519 VLMVPETKG 527
DB 418 IAIVPETKG 426

RESULT 10
AAG38870
ID AAG38870 standard; protein; 448 AA.
XX
AC AAG38870;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 48017.
XX

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.

PR	01-SEP-1999;	99US-01519303P;
PR	07-SEP-1999;	99US-01523633P;
PR	10-SEP-1999;	99US-01533070P;
PR	13-SEP-1999;	99US-01533758P;
PR	15-SEP-1999;	99US-01540181P;
PR	16-SEP-1999;	99US-01540393P;
PR	20-SEP-1999;	99US-01545779P;
PR	22-SEP-1999;	99US-01551139P;
PR	23-SEP-1999;	99US-01554866P;
PR	24-SEP-1999;	99US-01556591P;
PR	28-SEP-1999;	99US-01556458P;
PR	29-SEP-1999;	99US-01556596P;
PR	04-OCT-1999;	99US-01571117P;
PR	05-OCT-1999;	99US-01577533P;
PR	06-OCT-1999;	99US-01578653P;
PR	07-OCT-1999;	99US-01580291P;
PR	08-OCT-1999;	99US-01582322P;
PR	12-OCT-1999;	99US-01583692P;
PR	13-OCT-1999;	99US-01592932P;
PR	13-OCT-1999;	99US-01592942P;
PR	13-OCT-1999;	99US-01592952P;
PR	14-OCT-1999;	99US-01593292P;
PR	14-OCT-1999;	99US-01593303P;
PR	14-OCT-1999;	99US-01593331P;
PR	14-OCT-1999;	99US-01596372P;
PR	14-OCT-1999;	99US-01596382P;
PR	18-OCT-1999;	99US-01595842P;
PR	21-OCT-1999;	99US-01607412P;
PR	21-OCT-1999;	99US-01607672P;
PR	21-OCT-1999;	99US-01607682P;
PR	21-OCT-1999;	99US-01607702P;
PR	21-OCT-1999;	99US-01608142P;
PR	21-OCT-1999;	99US-01608152P;
PR	22-OCT-1999;	99US-01609380P;
PR	22-OCT-1999;	99US-01609381P;
PR	22-OCT-1999;	99US-01609389P;
PR	25-OCT-1999;	99US-01614042P;
PR	25-OCT-1999;	99US-01614052P;
PR	25-OCT-1999;	99US-01614062P;
PR	26-OCT-1999;	99US-01613592P;
PR	26-OCT-1999;	99US-01613602P;
PR	26-OCT-1999;	99US-01613612P;
PR	28-OCT-1999;	99US-01619202P;
PR	28-OCT-1999;	99US-01619922P;
PR	28-OCT-1999;	99US-01619932P;
PR	29-OCT-1999;	99US-01621422P;

Query Match 33.6%; Score 933; DB 3; Length 448;
Best Local Similarity 42.0%; Pred. No. 5.9e-78;
Matches 180; Conservative 102; Mismatches 141; Indels

QY	100	FLCTLLIVALGPIQFGFTSGFSSPTQDAMVRDLNLSISEFSAFGLSLNVGGMVGAIASGQM	159
Db	13	YLSFTFVAVCGSFAGSCAGYSSPAQAAIRNDLSLTIAEFSLFGSLTLTFGAMIGAITSGPI	72
QY	160	AEYIGRKGSMLTAAIPNIIGWLAIISFAKDA SFLYMGRLLLEGFGVGIIISYTVPVYIAEISP	219
Db	73	ADLVGRKGAMRVSSAFVCVVGWLAIIFAKGVVALDLGRLATGYGMGAFSYVVPFI FIAEIA P	132
QY	220	QNMRGALGSVN-QLSVTFGIFLAYLLGMFIPWRLLAGVICALPCTMLIPGLFFIPESPRWL	278
Db	133	KTFRGALTTLNQIILICTGVSVSFIIGTLVTRVRLALIGIIPCAASFLGLFFIPESPRWL	192
QY	279	AKMNLTEDCETSLQVLRGFETDITTEVNDIKRAVASSSKRTTISFQELNQKKYRTPLLLG	338
Db	193	AKVGRDTEFEALRKLRGKKADI SEEAEEIQDYIETLERLPKAKMLDLFQRRYIRSVLIA	252
QY	339	IGLLVLQNLSGINGVLFYASSIFPKAAGVTNSDLATCSLGAIQVLATGVTTWLLDRGRRI	398
Db	253	FGLMVFPQFGGINGICFYTSSIFEQAGFP-TRLGMIIVAVLQVVITALNAPIVDAGRKP	311
QY	399	LLIISTSGMTLCLLAVSVVFFFLKDNISQDSNSYYILTMISLVGIVSFVITFSFGMGAIPW	458

Db 312 LLLVSATGLVIGCLIAAASFYLVKH----DMAHEAVPVLAVVGIMVYIGSFSAGMGAMPW 367
QY 459 LMMSEILPVSIKSLGGSIAITLANWLTSAITMTTNTNMLTWSVGGTFLSYMVVSAFTIVFV 518
Db 368 VVMSEIFPINIKGVAGGMATLVNMFAGAWAVSYTFNFLMSWSSYGTFLIYAAINALAIVFV 427
QY 519 VLWVPETKG 527
Db 428 IAIVPETKG 436

RESULT 11
AAG38869
ID AAG38869 standard; protein; 464 AA.
XX
AC AAG38869;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 48016.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.

PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.

PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 33.6%; Score 933; DB 3; Length 464;
Best Local Similarity 42.0%; Pred. No. 6.2e-78;
Matches 180; Conservative 102; Mismatches 141; Indels 6; Gaps 3;

Qy 100 FLCTLIVLGP IQFTSGFSSPTQDAMVRDLNLSISEFSAFGSLSNVGMVGAIASGQM 159

Db 29 YLSTFVAVCGSFAFGSCAGYSSPAQAIRNDLSLTIAEFSLFGSLTFGAMIGAITSGPI 88
Qy 160 AEYIGRKGLMIAAIPNIIGWLAIISFAKDASFLYMGRLLEGFGVGIIISYTPVYIAEISP 219
Db 89 ADLVGRKGAMRVSSAFCVWGWLAIIFAKGVVALDLGLRLATGYGMAFSYVVPYIFIAEIA 148
Qy 220 QNMRGALGSVN-QLSVTFGIFLAYLLGMFIPWRLLAVIGALPCTMLIPGLFFIPESPRWL 278
Db 149 KTFRGALTTLNQILICTGVSVSFIIGTLVTRVRLALIGIIPCAASFLGLFFIPESPRWL 208
Qy 279 AKNLTEDCETSLQVLRGFETDITTEVNDIKRAVASSSKRTTISFQELNQKKYRTPLLLG 338
Db 209 AKVGRDTEFEAALRKLRGKKADISEEAAEIQDYIETLERLPKAKMLDLFQRRYIRSVLIA 268
Qy 339 IGLLVQLNSGINGVLFYASSIFKAAGVTNSDLATCSLGAIOVLATGVTTWLLDRAGRI 398
Db 269 FGLMVFPQQFGGINGICFYTSSIFEQAGFP-TRLGMIYAVLQVVTALNAPIVDRAGRKP 327
Qy 399 LLIISTSGMTCLLAVSVVFFLKDNIQSQNSYYILTMISLVGIVSFVITFSFGMGAIPW 458
Db 328 LLLVSATGLVIGCLIAAVSFYLVKH-----DMAHEAVPVLAVVGIMVVGISFSGMGMAMPW 383
Qy 459 LMMSEILPVSIXSLGGSIAATLANWLTSPAITMTNMLTWSVGGTFLSYMVVSFAFTIVFV 518
Db 384 VVMSEIFPINIKGVAGGMATLVNWFAGAVSVYTFNFLMSWSSYGTFLIYAAINALAIVFV 443
Qy 519 VLWVPETKG 527
Db 444 IAIVPETKG 452
RESULT 12
ADM48307
ID ADM48307 standard; protein; 481 AA.
XX AC ADM48307;
XX DT 03-JUN-2004 (first entry)
DE Polypeptide sequence #357 useful in producing transgenic plants.
XX KW Plant; maize; rice; sorghum; thale cress; soybean; wheat; transgenic;
KW osmotic stress; sugar transport; cell cycle pathway; plant height;
KW carbohydrate transport; crop productivity; plant growth;
KW stress resistance; disease resistance; insect resistance; heat tolerance;
KW nitrogen assimilation; water stress tolerance;
KW photosynthetic carbon fixation; virus resistance; gene therapy.
XX OS Glycine max.
XX PN US2003233670-A1.
XX PD 18-DEC-2003.
XX PF 04-DEC-2002; 2002US-00310154.
XX PR 04-DEC-2001; 2001US-0337358P.
XX PA (EDGE/) EDGERTON M D.
PA (CHOM/) CHOMET P S.
PA (LACC/) LACCETTI L B.
XX PI Edgerton MD, Chomet PS, Laccetti LB;
XX DR WPI; 2004-061374/06.
DR N-PSDB; ADM47939.
XX PT New polynucleotide, useful for manipulating plant protein quality,
PT improving plant growth, yield and crop productivity or grain composition
PT or producing plants with improved properties.
XX PS Claim 8; SEQ ID NO 725; 144pp; English.

XX CC The present invention relates to polynucleotide sequences, and the
CC proteins they encode. The sequences are isolated from a variety of
CC organisms such as plants (e.g. maize, rice, sorghum, thale cress,
CC soybean, and wheat), cyanobacteria, bacteria, yeast and other fungi. The
CC polynucleotide and polypeptide sequences of the invention are useful in
CC the production of transgenic plants that have improved properties. Also
CC disclosed are methods of producing fertile transgenic plants, preferably
CC maize, with desired phenotypes. The polynucleotide and polypeptide
CC sequences are useful for improving plants by providing protection against
CC osmotic stress, improving altering sugar transport and/or metabolism,
CC modifying the cell cycle pathway, reducing plant height, modifying
CC carbohydrate transport, improving crop productivity, improving plant
CC growth and stress resistance, improving disease resistance, improving
CC insect resistance, improving cold or heat tolerance, improving nitrogen
CC assimilation, improving stalk strength, improving water stress tolerance,
CC improving photosynthetic carbon fixation, improving biotic and abiotic
CC stress resistance, improving resistance to oxidative stress, providing
CC increased vigour, reducing senescence, and conferring virus resistance.
CC The present sequence represents a polypeptide sequence of the invention.
CC Note: The sequence data for this patent is not provided in the printed
CC specification but is obtained in electronic format from the USPTO website
CC at seqdata.uspto.gov.

Query Match	32.1%; Score 891; DB 8; Length 481;
Best Local Similarity	37.2%; Pred. No. 5.4e-74;
Matches 188; Conservative 97; Mismatches 171; Indels 50; Gaps 4;	
QY	48 AGAGEESGDHGVLRRLNLTGTSWYRMSSRQSGFAPGTSSMAVLRESHVSAFLCTLIVA 107
Db	25 AAACKENGSD-----KSVKNQSGIGMVL-----LSTLVAV 53
QY	108 LGPIQFGFTSGFSSPTQDAMVRDLNLSISEFSAFGSLSNVGMVGAIASGQMAEYIGRK 167
Db	54 CGSFTFGTCVGSAPTQAAIRADLNLSLAEFMFGSLVTIGAMLGAIITSGRITDFIGRK 113
QY	168 SLMIAAIPNIIGWLAIISPAKDAFLYMGRLLEGFGVGIISYTPVPIYIAEISPONMRGALG 227
Db	114 AMRISTGFCITGWIAVFFSKGSYSLDGFRFTGVIIGVISVVPVPIYIAEIAPIKNLRGLA 173
QY	228 SVNQLSVTFGIFLAYLLGMFIPWRLLAVIGALPCTMLIPGLFFIPESPRWLAKNLTEDC 287
Db	174 TTNQLLIVTGCSSVFLGSGVINWRELALAGLVPCICLLVGLCFIPESPRWLAKVGREKEF 233
QY	288 ETSLOVLRGFETDITTEVNDIKRAVASSSKRTTISFQELNQKRYRTPLLLGIGLLVLQNL 347
Db	234 QLALSRLRGKDADISDEAAEILDYIETLQSLPDKTLDLDFQSKYVHSVIVIGVLMACQQS 293
QY	348 SGINGVLFPYASSIFKAAGVTNSDLATCSLGAIQVLATGVTTWLLDRAGRRIILIISTSGM 407
Db	294 VGINGIGFYTAEIFVAAGLSSGKAGTIAVACIQIPFTLLGAILMDKSGRRPLVMVSAAGT 353
QY	408 TLCLLAVSVVFFLKDNIQSODSNSYYILTMISLVGIVSFVITFSFGMGAIPHLMVSEILPV 467
Db	354 FLGCFVAAFAFFLKD---QSLLPWVVPILAFAGVLIIYIAAFSIGLSGVPVIMSEIFPI 409
QY	468 STKSLGCSIATLANWLTSFAITWTTNMLTWVSGGTFLSYMVVSAAFTIVFVVLWVPETKG 527
Db	410 HLKGTAGSLVVLVAWLGAWVVSYTFNFMLNWSWSSPGTLFLYAGCSLLTILFAKLVPETKG 469
QY	528 XNSRGDTIFVLSLIQRQLQWLPECLS 553
Db	470 -----KTLSEIIQACIS 480

RESULT 13
AAG39282
ID AAG39282 standard; protein; 468 AA.

DT	18-OCT-2000	(first entry)
XX		
DE	Arabidopsis thaliana	protein fragment SEQ ID NO: 48581.
XX		
KW	Protein identification; signal transduction pathway; me	
KW	hybridisation assay; genetic mapping; gene expression c	
KW	termination sequence.	
XX		
OS	Arabidopsis thaliana.	
XX		
PN	EP1033405-A2.	
XX		
PD	06-SEP-2000.	
XX		
PF	25-FEB-2000; 2000EP-00301439.	
XX		
PR	25-FEB-1999; 99US-0121825P.	
PR	05-MAR-1999; 99US-0123180P.	
PR	09-MAR-1999; 99US-0123548P.	
PR	23-MAR-1999; 99US-0125788P.	
PR	25-MAR-1999; 99US-0126264P.	
PR	29-MAR-1999; 99US-0126785P.	
PR	01-APR-1999; 99US-0127462P.	
PR	06-APR-1999; 99US-0128234P.	
PR	08-APR-1999; 99US-0128714P.	
PR	16-APR-1999; 99US-0129845P.	
PR	19-APR-1999; 99US-0130077P.	
PR	21-APR-1999; 99US-0130449P.	
PR	23-APR-1999; 99US-0130510P.	
PR	23-APR-1999; 99US-0130891P.	
PR	28-APR-1999; 99US-0131449P.	
PR	30-APR-1999; 99US-0132048P.	
PR	30-APR-1999; 99US-0132407P.	
PR	04-MAY-1999; 99US-0132484P.	
PR	05-MAY-1999; 99US-0132485P.	
PR	06-MAY-1999; 99US-0132486P.	
PR	07-MAY-1999; 99US-0132487P.	
PR	07-MAY-1999; 99US-0132863P.	
PR	11-MAY-1999; 99US-0134256P.	
PR	14-MAY-1999; 99US-0134218P.	
PR	14-MAY-1999; 99US-0134219P.	
PR	14-MAY-1999; 99US-0134221P.	
PR	14-MAY-1999; 99US-0134370P.	
PR	18-MAY-1999; 99US-0134768P.	
PR	19-MAY-1999; 99US-0134941P.	
PR	20-MAY-1999; 99US-0135124P.	
PR	21-MAY-1999; 99US-0135353P.	
PR	24-MAY-1999; 99US-0135629P.	
PR	25-MAY-1999; 99US-0136021P.	
PR	27-MAY-1999; 99US-0136392P.	
PR	28-MAY-1999; 99US-0136782P.	
PR	01-JUN-1999; 99US-0137222P.	
PR	03-JUN-1999; 99US-0137528P.	
PR	04-JUN-1999; 99US-0137502P.	
PR	07-JUN-1999; 99US-0137724P.	
PR	08-JUN-1999; 99US-0138094P.	
PR	10-JUN-1999; 99US-0138540P.	
PR	10-JUN-1999; 99US-0138847P.	
PR	14-JUN-1999; 99US-0139119P.	
PR	16-JUN-1999; 99US-0139452P.	
PR	16-JUN-1999; 99US-0139453P.	
PR	17-JUN-1999; 99US-0139492P.	
PR	18-JUN-1999; 99US-0139454P.	
PR	18-JUN-1999; 99US-0139455P.	
PR	18-JUN-1999; 99US-0139460P.	
PR	18-JUN-1999; 99US-0139456P.	
PR	18-JUN-1999; 99US-0139457P.	
PR	18-JUN-1999; 99US-0139458P.	
PR	18-JUN-1999; 99US-0139459P.	
PR	18-JUN-1999; 99US-0139460P.	
PR	18-JUN-1999; 99US-0139461P.	
PR	18-JUN-1999; 99US-0139462P.	
PR	18-JUN-1999; 99US-0139463P.	
PR	18-JUN-1999; 99US-0139750P.	

Db 261 RYAPSVVIGVGLMLLQQLSGSSGLMYVVGSVFDKGGFPSS-IGSMILAVIMIPKALLGLI 311

QY 390 LLDRAGRRILLIISTSGMTLCLLAVSVVFFLKNISQDSNSYYIL- - - - - TMISLVGIVSF 445

Db 320 LVERKMGRRPULLASTGGMCFSSLLSFSFCFR- - - - - SYGMLDELTPIFTCIGVVG 371

QY 446 VITFSFGMGAIPWLMMSEILPVSIXSLGSGSIATLANWLTSPFAITMTTNLMLTWSVGGTFL 505

Db 372 ISSFAVGMGGLPWIIMSEIIFPMNVKVSAGTLVTLANWSFGWIVAFAYNFMLEWNASGTFL 431

QY 506 SYMVVSAFTIVFVVLWVPETKG 527

Db 432 IFFTICGAGIVFIYAMVPETKG 453

RESULT 14

AAG39281

ID AAG39281 standard; protein; 470 AA.

XX AC AAG39281;

XX 18-OCT-2000 (first entry)

DT DT

XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 48580.

XX

KW Protein identification; signal transduction pathway; metabolic pathway;

KW hybridisation assay; genetic mapping; gene expression control; promoter;

KW termination sequence.

XX

OS Arabidopsis thaliana.

XX

PN EP1033405-A2.

XX

PD 06-SEP-2000.

XX

PF 25-FEB-2000; 2000EP-00301439.

XX

PR 25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-0123180P.

PR 09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125788P.

PR 25-MAR-1999; 99US-0126264P.

PR 29-MAR-1999; 99US-0126785P.

PR 01-APR-1999; 99US-0127462P.

PR 06-APR-1999; 99US-0128234P.

PR 08-APR-1999; 99US-0128714P.

PR 16-APR-1999; 99US-0129845P.

PR 19-APR-1999; 99US-0130077P.

PR 21-APR-1999; 99US-0130449P.

PR 23-APR-1999; 99US-0130510P.

PR 28-APR-1999; 99US-0130891P.

PR 28-APR-1999; 99US-0131449P.

PR 30-APR-1999; 99US-0132048P.

PR 30-APR-1999; 99US-0132407P.

PR 04-MAY-1999; 99US-0132484P.

PR 05-MAY-1999; 99US-0132485P.

PR 06-MAY-1999; 99US-0132486P.

PR 06-MAY-1999; 99US-0132487P.

PR 07-MAY-1999; 99US-0132863P.

PR 11-MAY-1999; 99US-0134256P.

PR 14-MAY-1999; 99US-0134218P.

PR 14-MAY-1999; 99US-0134219P.

PR 14-MAY-1999; 99US-0134221P.

PR 14-MAY-1999; 99US-0134370P.

PR 18-MAY-1999; 99US-0134768P.

PR 19-MAY-1999; 99US-0134941P.

PR 20-MAY-1999; 99US-0135124P.

PR 21-MAY-1999; 99US-0135353P.

PR 24-MAY-1999; 99US-0135629P.

PR 25-MAY-1999; 99US-0136021P.

PR 27-MAY-1999; 99US-0136392P.

PR 28-MAY-1999; 99US-0136782P.

PR 01-JUN-1999; 99US-0137222P.

PR	05-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161922P.
PR	28-OCT-1999;	99US-0161993P.
PR	29-OCT-1999;	99US-0162142P.

Query Match

31.5%; Score 874.5; DB 3; Length 470;

Best Local Similarity 39.4%; Pred. NO. 1.8e-72;					
Matches 174; Conservative 95; Mismatches 158; Indels 15; Gaps 4;					
QY	92	LRESHVSA--FLCTLI	VALGPIQFGFTSGFSSPTQDAMVRDLNLSISEFSAFGLSNVGG 149		
Db	23	INECRITAVVLFSTFVS	CGSFCFGCAGYSSVAQTGIINDLGLSVAQYSMEFGSIMTFGG 82		
QY	150	MVGAIASGQMAEYIGRK	GSLMIAAIPNIIGWLAI	SAFKDASFLYMGRLLEGFGVGIISYT 209	
Db	83	MIGAI	FSGKVADLMGRKGTWFAQIFCIFGWAVA	LAKDSMWLDDIGRLSTGFAVGLLSYV 142	
QY	210	VPVYIAEISPNMRGALGS	VNQLSVTFGIFLAYLLGMFIPWRLLAVIGALPCTMLIPGLF 269		
Db	143	IPVYIAEITPKHVRGAFV	ANQLMQSCGLSLFYVIGNFVHWRNLALIGLIPCALQVVTLF 202		
QY	270	FIPESPRWLAKMNLTE	DCE	TSLOVLRGFETDITTEVNDIKRAVASSSKRTTISFQELNQK 329	
Db	203	FIPESPRLLGKWGHEKE	CRASLQSLRGDDADISEANTIKETMILFDEGPKSRVMDLFQR 262		
QY	330	KYRTP	LLLGIGLLVQLNSGINGVLFYASSIFKAAAGVTNSDLATCSLGAIQVLATGVTTW 389		
Db	263	RYAPSVVIGVGLMLLQ	LSGSSGLMYVGVSVFDKGGFPSS-IGSMILAVIMIPKALLGLI 321		
QY	390	LLDRAGRRILLIISTG	MTCLCLLAVSVVFFLKDNISQDSNSYYIL----	TMISLVGIVSF 445	
Db	322	LVEKMGRRP	LLLASTGGMCF	FFSLLLSFSCFR-----SYGMLDELTPIFTCIGVVG 373	
QY	446	VITFSFGMGAIPWLM	MEILPVS	IKSLGGS	IATLANWLTSFAITMTNMLTWSVGGTFL 505
Db	374	ISSFAVGMGGLPWIM	SEIFPMNVKVSAGTLVTLANWSFGWIVAFAYNFMLEWNASGTFL 433		
QY	506	SYMVVS	AFTIVFVVLWVPETKG 527		
Db	434	IFFTICGAGIVFI	YAMVPETKG 455		

RESULT 15

AAG50066

ID AAG50066 standard; protein; 474 AA.

XX AAG50066;

XX 18-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 63406.

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

OS Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-00301439.

XX 25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-0123180P.

PR 09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125788P.

PR 25-MAR-1999; 99US-0126264P.

PR 29-MAR-1999; 99US-0126785P.

PR 01-APR-1999; 99US-0127462P.

PR 06-APR-1999; 99US-0128234P.

PR 08-APR-1999; 99US-0128714P.

PR 16-APR-1999; 99US-0129845P.

PR 19-APR-1999; 99US-0130077P.

PR 21-APR-1999; 99US-0130449P.

PR 23-APR-1999; 99US-0130510P.

PR 23-APR-1999; 99US-0130891P.

PR 28-APR-1999; 99US-0131449P.

PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.

PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.

PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161922P.
PR	28-OCT-1999;	99US-0161993P.
PR	29-OCT-1999;	99US-0162142P.
Query Match 31.3%; Score 868.5; DB 3; Length 474;		
Best Local Similarity 41.1%; Pred. No. 6.6e-72;		
Matches 181; Conservative 94; Mismatches 150; Indels 15; Gaps 6;		
Qy	94	ESHVSA--FLCTLIJVALGPIQFGFTSGFSSPTQDAMVRDLNLSISEFSAFGSLSNVGGMV 151
Db	24	ECRITACVILSTFIAVCGSFSGVSLGYTSGAEIGIMKDLDSIAQFSAFASLSTLGAAI 83
Qy	152	GAIASGQMAEYIGRKGLMIAAIPNIIGWLAIISFAKDAFLYMGRLLEGFGVGIIISYVTP 211
Db	84	GALFSGKMAIILGRKTMWVSDLLCICIGWFESIAFAKDVWMLNPGRISSGIGLGLISYVVP 143
Qy	212	VYIAEISQNMREGALGSVNQLSVTFGIFLAYLLGMFIPWRLLAIVIGALPCTMLIPGLFFI 271
Db	144	VYIAEISPKHVRGTFFTNQLLQNSGLAMVYFSGNFNLRILALGALPCFIQVIGLEFV 203
Qy	272	PESPRWLAKMNLTEDCETSLQVLRGFETDITTEVNDIK---RAVASSSKRTTISFQELNQ 328
Db	204	PESPRWLAKVGSCKELENLLRLRGGNADISREASDIEVMTKMVENDSKS---SFCDLFQ 260
Qy	329	KKYRTPLLGIGLLVLQNLNGINGVLFYASSIFKAAAGVTNSDLATCSLGAIQVLATGVT 388
Db	261	RKYRYTLVVGIGLMLIQQFSGSSAVLSYASTILRKAGFSVT-IGSTLLGLFMIPKAMIGV 319
Qy	389	WLLDRAGRRILLIISTSGMTLCLLAVSVVFFL-KDNISQDSNSYYILTMISLVGIVSFVI 447
Db	320	ILVDKWGRRPRLLTSVSGMCITSMILIGVAFTLQKQMLLPETLPVFTFICVTL-----YIG 374
Qy	448	TFSFGMGAIPWLMNSEILPVSIKSLGGSATLANWLTSFAITMTNMLTWSVGGTFLSY 507
Db	375	TYAIGLGLPWIMSEIFPMNIKVTAGSIVTLVSWSSSSIVTYAFNLEWSTQGTFFVF 434
Qy	508	MWVSAFTIVFVVLWVPETKG 527
Db	435	GAVGGLALLFIWLLVPETKG 454

Search completed: October 13, 2004, 11:38:31
Job time : 60.5861 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2004, 11:33:23 ; Search time 17.1663 Seconds
(without alignments)
2136.387 Million cell updates/sec

Title: US-10-051-909-36

Perfect score: 2779
Sequence: 1 PSSSSSFRPAGKKKKKNQG.....TIFVSLSIQRQLWLPECLS 553

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	607	21.8	488	4	US-10-162-012-46 Sequence 46, Appl
2	575.5	20.7	524	2	US-08-928-692-12 Sequence 12, Appl
3	575.5	20.7	524	3	US-09-339-972-12 Sequence 12, Appl
4	572.5	20.6	494	2	US-09-031-392-5 Sequence 5, Appl
5	572.5	20.6	494	3	US-09-299-549-5 Sequence 5, Appl
6	572.5	20.6	494	3	US-09-610-417-5 Sequence 5, Appl
7	568	20.4	493	2	US-09-031-392-10 Sequence 10, Appl
8	568	20.4	493	3	US-09-299-549-10 Sequence 10, Appl
9	568	20.4	493	3	US-09-610-417-10 Sequence 10, Appl
10	557.5	20.1	539	3	US-09-291-922-26 Sequence 26, Appl
11	550	19.8	509	2	US-09-031-392-6 Sequence 6, Appl
12	550	19.8	509	3	US-09-299-549-6 Sequence 6, Appl
13	550	19.8	509	3	US-09-610-417-6 Sequence 6, Appl
14	550	19.8	510	3	US-09-291-922-22 Sequence 22, Appl
15	542.5	19.5	488	2	US-08-928-692-10 Sequence 10, Appl
16	542.5	19.5	488	3	US-09-339-972-10 Sequence 10, Appl
17	542.5	19.5	523	3	US-09-291-922-24 Sequence 24, Appl
18	527.5	19.0	513	3	US-09-291-922-20 Sequence 20, Appl
19	524	18.9	863	4	US-09-894-927B-9 Sequence 9, Appl
20	522	18.8	529	3	US-09-291-922-28 Sequence 28, Appl
21	522	18.8	549	3	US-09-291-922-30 Sequence 30, Appl
22	514.5	18.5	476	4	US-09-489-039A-11933 Sequence 11933, A
23	497	17.9	514	4	US-09-489-039A-11902 Sequence 11902, A
24	496.5	17.9	584	2	US-08-928-692-13 Sequence 13, Appl
25	496.5	17.9	584	3	US-09-339-972-13 Sequence 13, Appl
26	493.5	17.8	534	2	US-09-031-392-4 Sequence 4, Appl
27	493.5	17.8	534	3	US-09-299-549-4 Sequence 4, Appl

28	493.5	17.8	534	3	US-09-610-417-4 Sequence 4, Appl
29	486.5	17.5	492	2	US-08-355-844-3 Sequence 3, Appl
30	486.5	17.5	492	5	PCT-US95-16126-3 Sequence 3, Appl
31	471	16.9	514	4	US-09-679-686B-22 Sequence 22, Appl
32	470.5	16.9	500	2	US-09-031-392-7 Sequence 7, Appl
33	470.5	16.9	500	3	US-09-299-549-7 Sequence 7, Appl
34	470.5	16.9	500	3	US-09-610-417-7 Sequence 7, Appl
35	470.5	16.9	517	4	US-09-679-686B-18 Sequence 18, Appl
36	469	16.9	518	4	US-09-679-686B-23 Sequence 23, Appl
37	469	16.9	519	4	US-09-679-686B-24 Sequence 24, Appl
38	455	16.4	518	4	US-09-679-686B-16 Sequence 16, Appl
39	453.5	16.3	511	4	US-09-679-686B-12 Sequence 12, Appl
40	448	16.1	502	4	US-09-679-686B-2 Sequence 2, Appl
41	446	16.0	383	2	US-09-031-392-3 Sequence 3, Appl
42	446	16.0	383	3	US-09-299-549-3 Sequence 3, Appl
43	446	16.0	383	3	US-09-610-417-3 Sequence 3, Appl
44	446	16.0	501	4	US-09-489-039A-11731 Sequence 11731, A
45	442	15.9	552	4	US-09-248-796A-20713 Sequence 20713, A

ALIGNMENTS

RESULT 1
US-10-162-012-46
; Sequence 46, Application US/10162012
; Patent No. 6682597
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A.J.
; APPLICANT: Silos-Santiago, Inmaculada
; APPLICANT: Gu, Wei
; TITLE OF INVENTION: NOVEL HUMAN ION CHANNEL AND TRANSPORTER FAMILY MEMBERS
; FILE REFERENCE: 10448-190001
; CURRENT APPLICATION NUMBER: US/10/162,012
; CURRENT FILING DATE: 2002-06-04
; PRIOR APPLICATION NUMBER: US 60/209,845
; PRIOR FILING DATE: 2000-06-06
; PRIOR APPLICATION NUMBER: US 09/875,321
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: PCT/US01/18340
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/209,257
; PRIOR FILING DATE: 2000-06-05
; PRIOR APPLICATION NUMBER: US 09/875,423
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: PCT/US01/18398
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/209,238
; PRIOR FILING DATE: 2000-06-05
; PRIOR APPLICATION NUMBER: US 09/875,363
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: PCT/US01/18247
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/227,068
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: US 09/928,530
; PRIOR FILING DATE: 2001-08-13
; PRIOR APPLICATION NUMBER: PCT/US01/25475
; PRIOR FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: US 60/226,770
; PRIOR FILING DATE: 2000-08-21
; PRIOR APPLICATION NUMBER: US 09/934,421
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/26096
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/279,281
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 10/109,029
; PRIOR FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: PCT/US02/09728
; PRIOR FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: US 60/290,288
; PRIOR FILING DATE: 2001-05-11

; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/339,972
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA: 08/928,692
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 4944.200-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 524 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 6323002e
; US-09-339-972-12

Query Match 20.7%; Score 575.5; DB 3; Length 524;
Best Local Similarity 33.1%; Pred. No. 8.8e-47;
Matches 136; Conservative 94; Mismatches 147; Indels 34; Gaps 11;

QY 132 NLSISEFSAFGLSNVGGMVGAIASQMAEYIGRKGSMLMIAIPNIIGWLAIISFAK----D 188
Db 98 SLSVSSFA-----VGGMTASFFGGWLGDTLGRKAMLVANILSLVGLLGMGFSKLGPS 150

QY 189 ASFLYMGRLLEGFGVGIISYTPVYIAEISQPNMRGALGSVNQLSVTFGIF-----LAY 242
Db 151 HILIIAGRSISGLYGLISGLVPMYIGEIAPTALRGALGTFFHQLAIVTGTILISQIIIGLEF 210

QY 243 LLGMFIPWRLLAVIGALPCTMLIPGLFFIPESPRWL-AKMNLTEDCETSLQVLRGFETDI 301
Db 211 ILGNYDLWHILLGSGVRAILQSLLLFFCPESPRYLYIKLDEEVKAKQSLKRLRGYD-DV 269

QY 302 TTEVNDIK--RAVASSSKRTTISFQELNQKRYRTPLLLGIGLLVLQNLSGINGVLFYASS 359
Db 270 TKDINEMRKEREESSEQKVSIIQLFTNSSYRQPILVALMLHVAQQFSGINGIFYYSTS 328

QY 360 IFKAAGVNSDLATCSLGAIQVLATGVTWLLDRAGRRILLIISTSGMTLCLLAVSVVFF 419
Db 329 IFQTAGISKPVYATIGVAVNMVFTAVSVFLVEKAGRRSLFLIGMSGFMVCAIFMSVGLV 388

QY 420 LKDNISQDSNSYYILTMISLVGIVSFVITFSFGMGAIPWLMMSAILPVSIKSLGGSIAATL 479
Db 389 L-----LNKFSWMSYVSMIAIFLVSFFFGPGPIPFWMVAEFFSQGPRPAALAIAP 441

QY 480 ANWLTSPAITMTTNMLTWSVGGTFLSYM---VVSAPTIVFVVLVVPETKG 527
Db 442 SNWTCNFIVALCPQYIADFC--GPYVFFFLFAGVLLAFTL-FTFFKVPETKG 489

RESULT 4
US-09-031-392-5
; Sequence 5, Application US/09031392
; Patent No. 5942398
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; APPLICANT: Weng, Xun
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING GLUTEX AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street

; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/031,392
; FILING DATE: 26-FEB-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Meiklejohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 07334/072001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 494 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-031-392-5

Query Match 20.6%; Score 572.5; DB 2; Length 494;
Best Local Similarity 32.2%; Pred. No. 1.6e-46;
Matches 147; Conservative 86; Mismatches 173; Indels 51; Gaps 14;

QY 105 IVALGPIQFGFTSG-FSSPTQDAMVRD-LNLSISE-----FSAFCSLS----NVG 148
Db 16 IATIGSFQFGYNTGVINAP--EAIIKDFLNTLEERSETPPSSVLLTSLWSLSVAIFSVG 73

QY 149 GMVGAIASQMAEYIGRKGSMLMIAIPNIIGWLAIISPAKDA---SFLYMGRLLEGFGVGI 205
Db 74 GMIGFSVGLFVNRFRGRNSMLIVNLLAIAGCLMGFCCKIAESVEMLILGLLIGLFCGL 133

QY 206 ISYTPVYIAEISQPNMRGALGSVNQLSVTFGIFLAYLLGMFI-----PWRLLAIVIGAL 259
Db 134 CTGFVPMYIGEISPTALRGAFGTNLQNLGIVIGILVAQIFGLKVILGTEDLWPLLGLFTIL 193

QY 260 PCTMLIPGLFFIPESPRWLAKMNLTED--CETSLQVLRGFETDITTEVNDIKRAVASSSK 317
Db 194 PAIQCAALPFCPESPRELL-INRKEEEKAKEILLQRLWGTE-DVAQDIQEMKDESMRMSQ 251

QY 318 RTTISFQEL-NQKRYRTPLLLGIGLLVLQNLSGINGVLFYASSIFKAAGVTNSDLATCSL 376
Db 252 EKQVTVLELFRAPNYRQPIISIMLQSLQSLSGINAVFYYSTGIFKDAGVQEPVYATIGA 311

QY 377 GAIQVLATGVTWLLDRAGRRILLIISTSGMTLCLLAVSVVFFLKDNIQDSNSYYILTM 436
Db 312 GVVNTIFTVSVFLVERAGRRTLHLIGLGMAFCSILMTISLLLKDN-----YSWMSF 364

QY 437 ISLVGIVSFVITFSFGMGAIPWLMMSAILPVSIKSLGGSIAATLANWLTSPAITMTNML 496
Db 365 ICIGAILVFAFFGPIGPGPIPFWIVAEILFGQGPRAAMAVAGCSNWTSTNFLVG-----L 418

QY 497 TWSVGGTFLSYMVVSAPTIVFVVLW-----VPETKG 527
Db 419 LFPSATFYLGAYFIVFTVELVIFWVFTFFKVPETRG 455

RESULT 5
US-09-299-549-5
; Sequence 5, Application US/09299549
; Patent No. 6136547
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; APPLICANT: Weng, Xun
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING

Qy	314	SSSKT-----TISFOELNQKKYRTPLLLGIGLLVLQNLSGINGVLFYASSIFKAAAG	365
Db	286	SEEKRVLKDLILSPTIAMRHI-----LIAGIGIHFFQOSSGIDAVVLYSPLVFKSAG	337
Qy	366	VTNSDL---ATCSLGAIOVLATGVTTWLLDDRAGRRILLIISTSGMTLCLLAVSVVFFLKD	422
Db	338	ITGDSRLGTTAVAGATNTVFILVATFLLDRIRRRPLVLITSTGGMLVSLVGLATGLTVIS	397
Qy	423	NISQDSNSYYILTMISLVGIVSFVITFSFGMGAIPWLMMSBILPVSIKSLGGSIIATLANW	482
Db	398	RHPDEKITWAI--VLCIFCIMAAYVAFFSIGLGPITWTWYSSEIFPLHVRALGCSLGVAVNR	455
Qy	483	LTSFALTMT-TNLMLTWSVGGTFLSYMWVSAFTIVFVVLVWPETKG	527
Db	456	LTSGVISMTFISLSKAMTIGGAFELFAGIASFAWVFFFAVLPETRG	501

RESULT 11
US-09-031-392-6
; Sequence 6, Application US/09031392
; Patent No. 5942398
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; APPLICANT: Weng, Xun
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING GLUTEX AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
;

Query Match	19.8%;	Score 550;	DB 2;	Length 509;
Best Local Similarity	30.4%;	Pred. No. 2.5e-44;		
Matches 142;	Conservative 90;	Mismatches 177;	Indels 58;	Gaps 9;

QY	103	TLIVA-----LGPIQFGFTSGFSSPTQDAMVRDLN-----	132
		: : : : :	
db	23	TLVLA VFSAVLGLGLOFGYNIGVINAPOKVIEOSYNETWLGROGPEGPSSIPPGTLLTWLWA	82

QY 133 LSISEFSAFGSLNVGGMVCAIASGQMAEYIGRKGLMIAAIPNIGWLAI SFAKDAS-- 190

pb 83 LSVATES-----VGGMISSEFLIGTISOWIGRKRAMLVNNVLAVYIGSLMGLANAASY 135

```
QY      191 -FLYMGRLLGFGVCIISYTPVYIAEISPQNMRGALGSVNQLSVTFGIFLAY-----L 243
```

Dd	136	EMLI	GRFLIGAYSGLTSLG	VPMYGEIAPTHLR	GALGTNLQLAIVIGILIAQVLGLESL	195
Qy	244	LGMFIPWRILLAVIGALPCTMLIPGLFFIPESPRWL-AKMNLTEDCETSLQVLRGFE--TD	300			
Dd	196	LGTASLWPLLGLTVLPALLQLVLLPFCPESPRYLYIIQNLEGPARKSLKRLTGWADVSG	255			
Qy	301	ITTEVNDIKRAVASSSKRTTISFOELNQKKYRTPLLLGIGLLVLQNLSGINGLVFYASSI	360			
Dd	256	VLAELKDEKRKL--ERERPLSLLQLLGSRTHRQPLLIAVVQLSQQLSGINAVFYYSTSI	313			
Qy	361	FKAAGVTNSDLATCSLGAIOVLATGVTTWLLDRAGRRIILLISTSGMTCLLAVSVVFFL	420			
Dd	314	FETAGVGQPAYATIGAGVNVTVFTLVSVLLVERAGRRTLHLGLAGMCGCAILMTVALLL	373			
Qy	421	KDNISQDSNSYYILTMISLVGISVFVITFSFGMGAI PWLMVMSEILPVSIKSLGGSIATLA	480			
Dd	374	LERVP-----AMSVYSIVAIFGFVAFFEIGPGPIPFIVAELFSQGPRPAAMAVAGFS	426			
Qy	481	NWLTSFAITMTNMLTWSVGGTFLSYMWVSFTIVFVVLWVPETKG	527			
Dd	427	NWTSNFIIIGMFQYVAEAMGPYVFLLFVALLGLGFFIFTFLRVPETR	473			

RESULT 12
US-09-299-549-6
; Sequence 6, Application US/09299549
; Patent No. 6136547
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; APPLICANT: Weng, Xun
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING
; TITLE OF INVENTION: GLUTEX AND USES THEREOF
; NUMBER OF SEQUENCES: 10

Query Match 19.8%; Score 550; DB 3; Length 509;
Best Local Similarity 30.4%; Pred. No. 2.5e-44;
Matches 142. Conservative 90; Mismatches 177; Indels 58; Gaps 9;

```

Qy      103 TLIVA-----LGPIQFGFTSGFSSPTQDAMVRDLN-----132
      |||::|      |||::|::|      |      |      |      |

```

Db 23 TLVAVSAVLSLQFGYNIGVINAPQKVEIQSYNETWLGRCQPEGPSSIPPGTTLTTLWA 82
QY 133 LSISESAFGLSNVGGMVGAIASGQMAEYIGRKGLMTAAIPNIIGWLAIISFAKDAS-- 190
Db 83 LSVAFS-----VGGMISSFLIGIISQWLGRKRAMLVNNVLAIVLGGSLMGLANAAS 135
QY 191 -FLYMGRLLEGFGVGIISYTPVYIAEISPNMRGALGSVNQLSVTFGIFLAY-----L 243
Db 136 EMLILGRFLIGAYSGLTSGLVPMYVGEIAPTHLRGALGTNLQLAIVIGILIAQVLGLES 195
QY 244 LGMFIPWRLAVIGALPCTMLIPGLFFIPESPRWL-AKMNLTEDCETSLQVLRGFE--TD 300
Db 196 LGTASLWPLLLGLTVLPALLQLVLLPFCPESPRYLYIIQNLEGPARKSLKRLTGWADVSG 255
QY 301 ITTEVNDIKRAVASSSKRTTISFQELNQKRYRTPLLLGIGLLVLQNLSGINGVLFYASSI 360
Db 256 VLAEKDEKRL--ERERPLSLQLLGSRTTHRQPLIIAVVLQLSQQLSGINAVFYYSI 313
QY 361 FKAAGVNSDLATCSLGAIQVLATGVTWLLDRAGRRILLIISTSGMTLCLLAVSVVFFL 420
Db 314 FETAGVQPAYATIGAGVNTVFTLVSVLLVERAGRTHLLGLAGMCGCAILMTVALL 373
QY 421 KDNISQDSNSYYILTMISLVGIVSFVITFSFGMGAIPWLMSEILPVSIKSLGGSIA 480
Db 374 LERVP-----AMSVYSIVAIFGVFAFEIGPGPIPFIVAELEFSQGPRAAMAVAGFS 426
QY 481 NWLTSFAITMTNMLTWSVGGTFLSYMVVSFAFTIVFVVLWVPETKG 527
Db 427 NWTNSFIIGMGFYVAEAMGPVYVLLFAVLLGLGFFIFTLRVPETRG 473

RESULT 13

US-09-610-417-6
; Sequence 6, Application US/09610417
; Patent No. 6346374
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; Weng, Xun
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING
; GLUTEX AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C..
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/610,417
; FILING DATE: 05-Jul-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/299,549
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Meiklejohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 07334/072002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 509 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein

; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-610-417-6
Query Match 19.8%; Score 550; DB 3; Length 509;
Best Local Similarity 30.4%; Pred. No. 2.5e-44;
Matches 142; Conservative 90; Mismatches 177; Indels 58; Gaps 9;
QY 103 TLIVA-----LGPIQFGFTSGFSSPTQDAMVRDLN----- 132
Db 23 TLVAVSAVLSLQFGYNIGVINAPQKVEIQSYNETWLGRCQPEGPSSIPPGTTLTTLWA 82
QY 133 LSISESAFGLSNVGGMVGAIASGQMAEYIGRKGLMTAAIPNIIGWLAIISFAKDAS-- 190
Db 83 LSVAFS-----VGGMISSFLIGIISQWLGRKRAMLVNNVLAIVLGGSLMGLANAAS 135
QY 191 -FLYMGRLLEGFGVGIISYTPVYIAEISPNMRGALGSVNQLSVTFGIFLAY-----L 243
Db 136 EMLILGRFLIGAYSGLTSGLVPMYVGEIAPTHLRGALGTNLQLAIVIGILIAQVLGLES 195
QY 244 LGMFIPWRLAVIGALPCTMLIPGLFFIPESPRWL-AKMNLTEDCETSLQVLRGFE--TD 300
Db 196 LGTASLWPLLLGLTVLPALLQLVLLPFCPESPRYLYIIQNLEGPARKSLKRLTGWADVSG 255
QY 301 ITTEVNDIKRAVASSSKRTTISFQELNQKRYRTPLLLGIGLLVLQNLSGINGVLFYASSI 360
Db 256 VLAEKDEKRL--ERERPLSLQLLGSRTTHRQPLIIAVVLQLSQQLSGINAVFYYSI 313
QY 361 FKAAGVNSDLATCSLGAIQVLATGVTWLLDRAGRRILLIISTSGMTLCLLAVSVVFFL 420
Db 314 FETAGVQPAYATIGAGVNTVFTLVSVLLVERAGRTHLLGLAGMCGCAILMTVALL 373
QY 421 KDNISQDSNSYYILTMISLVGIVSFVITFSFGMGAIPWLMSEILPVSIKSLGGSIA 480
Db 374 LERVP-----AMSVYSIVAIFGVFAFEIGPGPIPFIVAELEFSQGPRAAMAVAGFS 426
QY 481 NWLTSFAITMTNMLTWSVGGTFLSYMVVSFAFTIVFVVLWVPETKG 527
Db 427 NWTNSFIIGMGFYVAEAMGPVYVLLFAVLLGLGFFIFTLRVPETRG 473

RESULT 14

US-09-291-922-22
; Sequence 22, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; -SEQ ID NO 22
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (102)
US-09-291-922-22
Query Match 19.8%; Score 550; DB 3; Length 510;
Best Local Similarity 29.6%; Pred. No. 2.5e-44;
Matches 141; Conservative 97; Mismatches 201; Indels 38; Gaps 8;
QY 99 AFLCTLIVAGLPIQFGFTSGFSSPTQDAMVRDLNLSISESAFGLSNVGGMVGAIASGQ 158
Db 21 AFACAILASMTSILLGYDIGVMGASLYIKDFNISDGKVEVLMGILNLYSLIGSFAAGR 80

THIS PAGE BLANK (USPIO)

OM protein - protein search, using sw model

Run on: October 13, 2004, 11:38:38 ; Search time 209.265 Seconds
(without alignments)
852.715 Million cell updates/sec

Title: US-10-051-909-36

Perfect score: 2779

Sequence: 1 PSSSSSFRPAGKKKKKQNG.....TIFVSLSTQRQLWLPECLS 553

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1356558 seqs, 322682953 residues

Total number of hits satisfying chosen parameters: 1356558

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2777	99.9	553	US-10-051-909-36	Sequence 36, Appl
2	2119.5	76.3	523	US-10-425-114-63431	Sequence 63431, A
3	2108.5	75.9	501	US-10-437-963-163544	Sequence 163544,
4	1801.5	64.8	501	US-10-051-909-38	Sequence 38, Appl
5	1796.5	64.6	509	US-10-425-114-59933	Sequence 59933, A
6	1718	61.8	484	US-10-424-599-275264	Sequence 275264,
7	1581	56.9	489	US-10-425-114-63430	Sequence 63430, A
8	1378	49.6	414	US-10-437-963-165948	Sequence 165948,
9	1219	43.9	318	US-10-425-114-65425	Sequence 65425, A
10	1199	43.1	361	US-10-424-599-275268	Sequence 275268,
11	1185	42.6	338	US-10-424-599-273908	Sequence 273908,
12	1032	37.1	219	US-10-425-114-49786	Sequence 49786, A
13	909.5	32.7	502	US-10-437-963-166774	Sequence 166774,
14	891	32.1	481	US-10-310-154-725	Sequence 725, App

15	768	27.6	216	15	US-10-424-599-275267	Sequence 275267,
16	759	27.3	544	16	US-10-437-963-106607	Sequence 106607,
17	758.5	27.3	263	15	US-10-424-599-275265	Sequence 275265,
18	717.5	25.8	421	15	US-10-424-599-247207	Sequence 247207,
19	658.5	23.7	477	14	US-10-168-651-1	Sequence 1, Appli
20	657.5	23.7	477	9	US-09-886-954-1	Sequence 1, Appli
21	625.5	22.5	228	15	US-10-425-114-47622	Sequence 47622, A
22	623	22.4	432	15	US-10-029-020-186	Sequence 186, App
23	618	22.2	507	14	US-10-157-031-359	Sequence 359, App
24	618	22.2	507	16	US-10-755-889-110	Sequence 110, App
25	612.5	22.0	487	9	US-09-795-693-27	Sequence 27, Appl
26	612.5	22.0	487	13	US-10-095-139-14	Sequence 14, Appl
27	612.5	22.0	487	14	US-10-156-239-27	Sequence 27, Appl
28	612.5	22.0	487	14	US-10-199-485-27	Sequence 27, Appl
29	612.5	22.0	487	14	US-10-391-399-45	Sequence 45, Appl
30	609	21.9	486	9	US-09-860-232A-7	Sequence 7, Appli
31	607	21.8	488	10	US-09-794-822-11	Sequence 11, Appl
32	607	21.8	488	13	US-10-094-059-4	Sequence 4, Appli
33	607	21.8	488	14	US-10-170-528-5	Sequence 5, Appli
34	607	21.8	488	14	US-10-162-012-46	Sequence 46, Appl
35	607	21.8	488	14	US-10-062-960B-4	Sequence 4, Appli
36	607	21.8	488	14	US-10-144-624-4	Sequence 4, Appli
37	607	21.8	488	14	US-10-162-102-46	Sequence 46, Appl
38	607	21.8	488	15	US-10-038-854-170	Sequence 170, App
39	605.5	21.8	480	9	US-09-734-569-172	Sequence 172, App
40	601.5	21.6	157	16	US-10-767-701-50013	Sequence 50013, A
41	600.5	21.6	191	15	US-10-425-114-61925	Sequence 61925, A
42	599	21.6	161	15	US-10-425-114-49791	Sequence 49791, A
43	597.5	21.5	389	16	US-10-767-701-44031	Sequence 44031, A
44	592.5	21.3	535	9	US-09-795-693-20	Sequence 20, Appl
45	592.5	21.3	535	14	US-10-156-239-20	Sequence 20, Appl

ALIGNMENTS

RESULT 1
US-10-051-909-36
; Sequence 36, Application US/10051909
; Publication No. US20020199217A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Helentjaris, Tim
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB1163 US CIP
; CURRENT APPLICATION NUMBER: US/10/051,909
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 60/083,044
; PRIOR FILING DATE: April 24, 1998
; NUMBER-OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 36
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (528)
; OTHER INFORMATION: Xaa = ANY AMINO ACID
US-10-051-909-36

Query Match 99.9%; Score 2777; DB 13; Length 553;
Best Local Similarity 100.0%; Pred. No. 1.5e-232;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PSSSSSFRPAGKKKKKQNGLRREAVPGRPASELRTRVMGGGSGNRGGAGAGEESGSDHDG 60

Db 1 PSSSSSFRPAGKKKKKQNGLRREAVPGRPASELRTRVMGGGSGNRGGAGAGEESGSDHDG 60

QY 61 VLRRPLLTGWSYRMSSRQSFAPGTSSMAVLRESHVSAFLCTLVALGPIQFGFTSCFS 120

Db 61 VLRRPLLTGSMYRMSSRQSSFAFGTSSMAVLRSHVSAFLCTLI VALGPIQFGFTSGFS 120

Qy 121 SPTQDAMVRDLNLSISEFSAFGLSNVGGMVGAIASGQMAEYIGRKGLMIAAIPNIGW 180

Db 121 SPTQDAMVRDLNLSISEFSAFGLSNVGGMVGAIASGQMAEYIGRKGLMIAAIPNIGW 180

Qy 181 LAISFAKDASFLYMGRLLEGGFVGVIISYTVPVYIAEISPNQMRGALGSVNQLSVTFGIFL 240

Db 181 LAISFAKDASFLYMGRLLEGGFVGVIISYTVPVYIAEISPNQMRGALGSVNQLSVTFGIFL 240

Qy 241 AYLLGMFIPWRLLAVIGALPCTMLIPGLFFIPESPRWLAKMNLTEDCETSLQVLRGFETD 300

Db 241 AYLLGMFIPWRLLAVIGALPCTMLIPGLFFIPESPRWLAKMNLTEDCETSLQVLRGFETD 300

Qy 301 ITTEVNDIKRAVASSSKRTTISFQELNQKKYRTPLLLGIGLLVLQNLSGINGVLFYASSI 360

Db 301 ITTEVNDIKRAVASSSKRTTISFQELNQKKYRTPLLLGIGLLVLQNLSGINGVLFYASSI 360

Qy 361 FKAAGVTNSDLATCSLGAIQVLATGVTTWLLDRAGRRILLIISTSGMTLCLLAVSVVFFL 420

Db 361 FKAAGVTNSDLATCSLGAIQVLATGVTTWLLDRAGRRILLIISTSGMTLCLLAVSVVFFL 420

Qy 421 KDNISQDSNSYYILTMISLVGIVSFVITFSFGMGAIPLWMMSEILPVSIKSLGGSIAATLA 480

Db 421 KDNISQDSNSYYILTMISLVGIVSFVITFSFGMGAIPLWMMSEILPVSIKSLGGSIAATLA 480

Qy 481 NWLTSFAITMTNLMLTWSVGGTFLSYMVVSAFTIVFVVLWVPETKGXNSRGDTIFVSLS 540

Db 481 NWLTSFAITMTNLMLTWSVGGTFLSYMVVSAFTIVFVVLWVPETKGXNSRGDTIFVSLS 540

Qy 541 IQRQLQWLPECLS 553

Db 541 IQRQLQWLPECLS 553

RESULT 2

US-10-425-114-63431

; Sequence 63431, Application US/10425114

; Publication No. US20040034888A1

; GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong

; APPLICANT: Zhou, Yihua

; APPLICANT: Kovalic, David K.

; APPLICANT: Screen, Steven E

; APPLICANT: Tabaska, Jack E

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53313)B

; CURRENT APPLICATION NUMBER: US/10/425,114

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 73128

; SEQ ID NO 63431

; LENGTH: 523

; TYPE: PRT

; ORGANISM: Oryza sativa

; FEATURE:

; OTHER INFORMATION: Clone ID: LIB4371-017-F11_FLI.pep

US-10-425-114-63431

Query Match 76.3%; Score 2119.5; DB 15; Length 523;

Best Local Similarity 81.7%; Pred. No. 2.1e-175;

Matches 419; Conservative 41; Mismatches 36; Indels 17; Gaps 5;

Qy 40 GGGSNR-GGAGAGEESGDHD--GVLRRPLL-NTGSWYRMSSRQSSF-APGTSSMAVLR 94

Db 20 GGAMNRGGVGVDSESGDYSGGMRKPLLMHTGSWYRMGRSQSLTGACTSSMAILRE 79

Qy 95 SHVSAFLCTLI VALGPIQFGFTSGFSSPTQDAMVRDLNLSISEFSAFGLSNVGGMVGA 154

Db 80 SHVSAFLCTLI VALGPIQFGFTSGFSSPTQDAIRDLDTLSEFSVFGSLSNVGAMVGA 139

Qy 155 ASGQMAEYIGRKGLMIAAIPNIIIGWLAIISFAKDASFLYMGRLLEGGFVGVIISYTVPVYI 214

Db 140 ASGQMAEYIGRKGLMIAAIPNIIIGWLAIISFAKDSSFLYMGRLLEGGFVGVIISYTVPVYI 199

Qy 215 AEISPNQMRGALGSVNQLSVTFGIFLAYLLGMFIPWRLLAVIGALPCTMLIPGLFFIPES 274

Db 200 AEISPNQMRGALGSVNQLSVTFGIFLAYLLGMFIPWRLLAVIGILPCTVLIPGLFFIPES 259

Qy 275 PRWLAKMNLTEDCETSLQVLRGFETDITTEVNDIKRAVASSSKRTTISFQELNQKKYRTP 334

Db 260 PRWLAKMNMDDFETSLQVLRGFETDISAEVNDIKRAVASANKRTTIRFQELNQKKYRTP 319

Qy 335 LLLGIGLLVLQNLSGINGVLFYASSIFKAAGVTNSDLATCSLGAIQVLATGVTTWLLDRA 394

Db 320 LILGIGLLVLQQLSGINGILFYAGSIFKAAGLTNSDLATCALGAIQVLATGVTTWLLDRA 379

Qy 395 GRRILLIISTSGMTLCLLAVSVVFFLKDNIQSODSNSYYILTMISLVGIVSFVITFSFGMG 454

Db 380 GRRILLIISAGMTLSLLAVAVVFFLKDNIQSODSHMYTSLMISLVALVAFVIAESFGMG 439

Qy 455 AIPWLMSEILPVSIKSLGGSIAATLIANWLTSFAITMTNLMLTWSVGGTFLSYMVVSAFT 514

Db 440 AIPWIINSEILPVSIKSLAGSFATLIANWLTSFGITMTANMLMWSAGGTFVSYMVVSAFT 499

Qy 515 IVFVVLWVPETKGXNSRGDTIFVLSIQRLQW 547

Db 500 LVFVILWVPETKGR-----LEEIQW 520

RESULT 3

US-10-437-963-163544

; Sequence 163544, Application US/10437963

; Publication No. US20040123343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 163544

; LENGTH: 501

; TYPE: PRT

; ORGANISM: Oryza sativa

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT4530_62530C.1.pep

US-10-437-963-163544

Query Match 75.9%; Score 2108.5; DB 16; Length 501;

Best Local Similarity 81.9%; Pred. No. 1.8e-174;

Matches 417; Conservative 41; Mismatches 34; Indels 17; Gaps 5;

Qy 44 NR-GGAGAGEESGDHD--GVLRRPLL-NTGSWYRMSSRQSSF-APGTSSMAVLRSHVS 98

Db 2 NRGGGVGVDSESGDYSGGMRKPLLMHTGSWYRMGRSQSLTGAGISSMAILRESHVS 61

Qy 99 AFLCTLI VALGPIQFGFTSGFSSPTQDAMVRDLNLSISEFSAFGLSNVGGMVGAISGQ 158

Db 62 AFLCTLI VALGPIQFGFTSGFSSPTQDAIIRDLDTLSEFSVFGSLSNVGAMVGAISGQ 121

Qy 159 MAEYIGRKGLMIAAIPNIIIGWLAIISFAKDASFLYMGRLLEGGFVGVIISYTVPVYIAEIS 218

Db 122 MAEYIGRKGLMIAAIPNIIIGWLAIISFAKDSSFLYMGRLLEGGFVGVIISYTVPVYIAEIS 181

Qy 219 PQNMRGALGSVNQLSVTFGIFLAYLLGMFIPWRLLAVIGALPCTMLIPGLFFIPESPRWL 278

Db 182 PNMRGALGSVNQLSVTVGILLAYLLGMFVWPWRLAVIGILPCTVLIPGLFFIPESRWL 241

QY 279 AKMNLTEDCETSLQVLRGFETDITTEVNDIKRAVASSSKRTTISFQELNQKKYRTPLLLG 338

Db 242 AKMNMDDFETSLQVLRGFETDISAEVNDIKRAVASANKRTTIRFQELNQKKYRTPLLLG 301

QY 339 IGLLVQLNLSGINGVLFFYASSIFKAAGVNDSLATCSLGAIQVLATGVTWLLDRGRRI 398

Db 302 IGLLVQLLSGINGILFYAGSIFKAAGLTNSDLATCALGAIQVLATGVTWLLDRGRRI 361

QY 399 LLIISTGMTCLLAVSVVFFLKDNISQDSNSYYILTMISLVGIVSFVITFSFGMGAIPW 458

Db 362 LLIISAGMTSLAVAVVFFLKDSISQDSHMYTTLMSISLVAVAFVIAFSFGMGAIPW 421

QY 459 LMMSEILPVSISKLSGSIATLANWLTSFAITMTNMLTWSVGGTFLSYMVVSAPTIVFV 518

Db 422 IIMSEILPVSISKLSAGSFATLANWLTSFGITMTANMLWSAGGTFFVSVMVSAFTLVFV 481

QY 519 VLWVPETKGXNSRGDTIFVSLSIQRQLQW 547

Db 482 ILWVPETKGR-----LEEIQW 498

RESULT 4

US-10-051-909-38

; Sequence 38, Application US/10051909

; Publication No. US20020199217A1

; GENERAL INFORMATION:

; APPLICANT: Allen, Steve

; APPLICANT: Helentjaris, Tim

; APPLICANT: Hitz, Bill

; APPLICANT: Kinney, Tony

; APPLICANT: Tingey, Scott

; TITLE OF INVENTION: Plant Sugar Transport Proteins

; FILE REFERENCE: BB1163 US CIP

; CURRENT APPLICATION NUMBER: US/10/051,909

; CURRENT FILING DATE: 2002-01-17

; PRIOR APPLICATION NUMBER: 60/083,044

; PRIOR FILING DATE: April 24, 1998

; NUMBER OF SEQ ID NOS: 38

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 38

; LENGTH: 501

; TYPE: PRT

; ORGANISM: Oryza sativa

US-10-051-909-38

Query Match 64.8%; Score 1801.5; DB 13; Length 501;

Best Local Similarity 72.2%; Pred. No. 8.8e-148;

Matches 351; Conservative 62; Mismatches 64; Indels 9; Gaps 2;

QY 51 GEESGSDHDG-----VLRRLNTGSWYRMSSRQSSFAPGT---SSMAVLRESHVSAFL 101

Db 5 GEESGGEDGGRTASASDLRKPFLLHTGSWYKMSAGCGGGMGSRGLGSSAYSRLDSSVSAVL 64

QY 102 CTLIVALGPIQFGFTSGFSSPTQDAMVRDLNLSISEFSAFGSLSNVGMVGAIASGQMAE 161

Db 65 CTLIVALGPIQFGFTCGFSSPTQDAIISDLGLTSLSEFSLFGSLSNVGMVGAIASGQIAE 124

QY 162 YIGRKGSMLIAAIPNIIIGWLAIISFAKDASFLYMGRLLEGFVGVIISYTPVYVIAEISPQN 221

Db 125 YIGRKGSMLIAAIPNIIIGWLAIISFAKDSSFLFMGRLLLEGFVGVIISYTPVYVIAEIPQT 184

QY 222 MRGALGSVNQLSVTFGIFLAVLLGMFIPWRLAVIGALPCTMLIPGLFFIPESPRWLAKM 281

Db 185 MRGALGSVNQLSVTIGILLAYLLGMFVPWRIILSVLGLIPCSILIPGLFFIPESPRWLAKM 244

QY 282 NLTEDCETSLQVLRGFETDITTEVNDIKRAVASSSKRTTISFQELNQKKYRTPLLLGIGL 341

Db 245 GKMEDFESSLQVLRGFETDIAVEVNEIKRSVQSSRRRTTIRFADIKQKRYSVPLMVGIGL 304

QY 342 LVLQNLSGINGVLFFYASSIFKAAGVNDSLATCSLGAIQVLATGVTWLLDRAGRRI 401

Db 305 LVLQQLSGVNGILFYAASIFKAAGLTNSNLATFGLGVVQVATGVTWLTDKAGRRLLLI 364

QY 402 ISTSGMTCLLAVSVVFFLKDNISQDSNSYYILTMISLVGIVSFVITFSFGMGAIPWLM 461

Db 365 ISTTGMTITLVVSVSFFVKDNITNGSHLYSVMSMLSLVGLVAFVIFSGLGALPWIIM 424

QY 462 SEILPVSISKLSGSIATLANWLTSFAITMTNMLTWSVGGTFLSYMVVSAPTIVFVVLW 521

Db 425 SEILPVSISKLSAGSVATLANWLTAWLITMTASMLWSWNGGTFAIYAAVCAGTLLVFVCLW 484

QY 522 VPETKG 527

Db 485 VPETKG 490

RESULT 5

US-10-425-114-59933

; Sequence 59933, Application US/10425114

; Publication No. US2004003488A1

; GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong

; APPLICANT: Zhou, Yihua

; APPLICANT: Kovalic, David K.

; APPLICANT: Screen, Steven E

; APPLICANT: Tabaska, Jack E

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53313)B

; CURRENT APPLICATION NUMBER: US/10/425,114

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 73128

; SEQ ID NO 59933

; LENGTH: 509

; TYPE: PRT

; ORGANISM: Zea mays

; FEATURE:

; OTHER INFORMATION: Clone ID: LIB3689-227-G11_FLI.pep

US-10-425-114-59933

Query Match 64.6%; Score 1796.5; DB 15; Length 509;

Best Local Similarity 70.4%; Pred. No. 2.4e-147;

Matches 352; Conservative 63; Mismatches 68; Indels 17; Gaps 4;

QY 41 GGSNRGGAGAGEESG-----SDHDGVLRRPLLNTGSWYR-----MSSRQSSFAPGTSS 88

Db 3 GMSFRDQESGGEDGGRTSSASD---LRKPLINTGSWYRMPPAGVMGSRQSSLMERLGS 58

QY 89 MAV-LRESHVSAFLCTLIVALGPIQFGFTSGFSSPTQDAMVRDLNLSISEFSAFGSLSNV 147

Db 59 SAFSLRDVAISATLCTLIVALGPIQFGFTCGYSSPTQDAIADLGLSLSEFSLFGSLSNV 118

QY 148 GGMVGAIASGQMAEYIGRKGSMLIAAIPNIIIGWLAIISFAKDASFLYMGRLLEGFVGVIIS 207

Db 119 GAMVGAISSGQLAEYIGRKGSMLIAAIPNIIIGWLAIISFAKDSSFLFMGRLLLEGFVGVIIS 178

QY 208 YTVPVYVIAEISPQNMRRGALGSVNQLSVTFGIFLAVLLGMFIPWRLAVIGALPCTMLIPG 267

Db 179 YTVPVYVIAEIPADQDORGALGSVNQLSVTIGILLAYLFGMFPWRIILAVLGLIPCSILIPG 238

QY 268 LFFIPESPRWLAKMNLTEDCETSLQVLRGFETDITTEVNDIKRAVASSSKRTTISFQELN 327

Db 239 LFFVPESPRWLAKMGKMEDEFYSLQVLRGFQTDITAEVNEIKRSLASSRRRTTIRFADIK 298

QY 328 QKKYRTPLLLGIGLLVLQNLSGINGVLFFYASSIFKAAGVNDSLATCSLGAIQVLATGVT 387

Db 299 QKRYSVPLVIGIGLLVLQQLSGVNGILFYAASIFKAAGITNSNLATFGLGAVQVIATGVT 358

QY 388 TWLLDRAGRRIILIIISTGMTCLLAVSVVFFLKDNISQDSNSYYILTMISLVGIVSFVI 447

Db 359 TWLTDKAGRRIILIIISTTGMVITLVIVSVSFFVKDNIAAGSHLYSVMSMLSLAGLVAFVI 418

[illegible]

RESULT 6
US-10-424-599-275264
; Sequence 275264, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 275264
; LENGTH: 484
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_90584C.1.pep
US-10-424-599-275264

Query Match	61.8%	Score 1718;	DB 15;	Length 484;
Best Local Similarity	69.5%	Pred. No. 1.5e-140;		
Matches 331; Conservative	71;	Mismatches 70;	Indels 4;	Gaps 2

Qy	52	EESGDH	DGVLR	RRP	LLNT	GT	SWY	R	M	S	S	R	Q	S	S	F	A	P	G	T	S	S	M	A	V	L	R	E	S	H	V	A	F	L	C	T	L	I	V	A	L	G	P	I	111																
Db	2	EESGD	A	-	-	L	Q	P	F	L	H	T	G	S	W	Y	K	M	G	S	R	Q	S	S	I	M	-	-	G	S	T	H	V	I	R	D	G	A	V	S	V	L	F	C	V	L	I	V	A	L	G	P	I	57							
Qy	112	QF	G	T	S	G	F	S	S	P	T	Q	D	A	M	V	R	D	L	N	L	S	I	S	E	F	S	A	F	G	S	L	S	N	V	G	M	V	G	A	I	A	S	G	M	A	E	Y	I	G	R	K	G	S	L	M	I	171			
Db	58	Q	F	G	T	C	G	Y	S	S	P	T	Q	G	A	I	V	R	D	L	N	L	S	I	S	E	F	S	F	F	G	S	L	S	N	V	G	A	M	V	G	A	I	A	S	G	I	A	E	Y	I	G	R	K	G	S	L	M	I	117	
Qy	172	A	A	I	P	N	I	I	G	W	L	A	I	S	F	A	K	D	A	S	F	L	Y	M	G	R	L	L	E	G	F	G	V	G	I	I	S	Y	T	P	V	Y	I	A	E	I	S	P	O	N	R	G	A	L	G	S	V	N	Q	231	
Db	118	A	A	I	P	N	I	I	G	W	L	A	I	S	F	A	K	D	S	F	L	Y	M	G	R	L	L	E	G	F	G	V	G	I	I	S	Y	V	P	V	Y	I	A	E	I	A	P	O	N	L	R	G	G	L	G	S	V	N	Q	177	
Qy	232	L	S	V	T	F	G	I	F	L	A	Y	L	L	G	M	F	I	P	W	R	L	L	A	V	I	G	A	L	P	C	T	M	L	I	P	G	L	F	F	I	P	E	S	P	R	W	L	A	K	N	L	T	E	D	C	E	T	S	L	291
Db	178	L	S	V	T	I	G	I	M	L	A	Y	L	L	G	F	V	N	W	R	V	L	A	I	L	G	I	L	P	C	T	V	L	I	P	G	L	F	F	I	P	E	S	P	R	W	L	A	K	M	I	D	E	F	E	T	S	L	237		
Qy	292	Q	V	L	R	G	P	E	T	D	I	T	T	E	V	N	D	I	K	R	A	V	A	S	S	K	R	T	T	I	S	F	O	E	L	N	Q	K	Y	R	T	P	L	L	L	G	I	L	L	V	L	O	N	L	S	G	I	N	351		
Db	238	Q	V	L	R	G	E	D	T	D	I	S	V	E	V	H	E	I	K	R	S	V	A	S	T	G	K	R	A	I	R	E	A	D	L	K	R	K	Y	W	F	P	L	M	V	G	I	L	L	V	L	Q	O	L	S	G	I	N	297		
Qy	352	G	V	L	F	Y	A	S	S	I	F	K	A	G	V	T	N	S	D	L	A	T	C	S	L	G	A	I	Q	V	L	A	T	G	V	T	T	W	L	L	D	R	A	G	R	R	I	L	I	I	S	T	S	G	M	T	L	C	L	411	
Db	298	G	I	L	F	Y	S	T	T	I	F	A	N	A	G	I	S	S	E	A	A	T	V	G	L	G	A	V	Q	V	I	A	T	G	I	S	T	W	L	V	D	K	S	G	R	R	L	L	I	I	S	S	S	V	M	T	V	S	L	357	
Qy	412	L	A	V	S	V	V	F	F	L	K	D	N	I	S	Q	D	S	N	S	Y	I	L	T	M	I	S	L	G	I	V	S	F	V	I	T	F	S	F	G	M	G	A	I	P	N	L	M	M	S	E	I	L	P	V	S	I	K	S	471	
Db	358	L	I	V	S	I	A	F																																																					

RESULT 7
US-10-425-114-63430
; Sequence 63430, Application US/10425114
; Publication No. US20040034888A1

```

; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 63430
; LENGTH: 489
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4309-039-E8_FLI.pep
US-10-425-114-63430

```

Query Match	56.9%;	Score 1581;	DB 15;	Length 489;
Best Local Similarity	82.8%;	Pred. NO. 1.3e-128;		
Matches 308;	Conservative	29;	Mismatches 23;	Indels 12;
				Gaps 1;

QY	176	NIIGWLAI	SFAKDASFLYMGRLL	EGVGIIISYTVPVYIAEISPQNMRGALGSVNQLSVT	235
		:	:	:	
Db	127	NIIGWLAI	SFAKDSFLYMGRLL	EGVGVIISYTVPVYIAEISPQNMRGALGSVNQLSVT	186
		:	:	:	
QY	236	FGIFLAYL	LGMFIPWRLLAVI	GALPCTMLIPGLFFIPESPRWLAKMNLTEDCETSLQVL	295
		:	:	:	
Db	187	VGILLAYL	LGMFVPWRLLAVI	GILPCITVLIPGLFFIPESPRWLAKNMMDDFETSLQVL	246
		:	:	:	
QY	296	GFETDITTE	VNDIKRAVASSKR	TTFISFOELNQKKYRTPLLLIGIGLLVLQNLSGINGVL	355
		:	:	:	
Db	247	GFETDISAE	VNDIKRAVASANKR	TTIRFOELNQKKYRTPILIGIGLLVLQQLSGINGIL	306
		:	:	:	
QY	356	YASSIFKRA	GVNTSDLATCSL	GAIQVLATGVTTWLLDRAGRRIILLIISTSGMTLCULLAVS	415
		:	:	:	
Db	307	YAGSIFKRA	GLTNSDLATCAL	GAIQVLATGVTTWLLDRAGRRIILLIISSAGMTLSLLAVA	366
		:	:	:	
QY	416	VVFELKDN	ISQDSNSYYILT	MISLVGIVSFVITFSFGMGAI PWLMMS EILPVSIKS LGS	475
		:	:	:	
Db	367	VVFELKDS	ISQDSHMYITLS	MISLVAFVIAFSGMGAI PWIIMSEILPVSIKSLAGS	426
		:	:	:	
QY	476	IATLANWL	TSPAITMTTNL	MLTWVGGTELSMVVSAPTIVFVVLWVPETKGNSRGDTI	535
		:	:	:	
Db	427	FATLANWL	TSGITMTANL	MLSWSAGGTIVFSVMVVSFTLVFVILWVPETKGR	480
		:	:	:	
QY	536	FVSLSIQR	QLQW	547	
		: :	:	:	
Db	481	-----LEEI	QW	486	

```

RESULT 8
US-10-437-963-165948
; Sequence 165948, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid M
; TITLE OF INVENTION: Plants and Uses Th
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,9
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966

```



```
; SEQ ID NO 165948
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(414)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_64704C.1.pep
US-10-437-963-165948

Query Match      49.6%; Score 1378; DB 16; Length 414;
Best Local Similarity 59.5%; Pred. No. 4.5e-111;
Matches 289; Conservative 46; Mismatches 55; Indels 96; Gaps 6;

QY 51 GEESGSDHDG-----VLRRPLNTGSMYRMSSRQSSFPAGT---SSMAVLRESHVSAFL 101
Db 5 GEESGGEDGGRTASADLRKPFLHTGSMYKMSAGGGGGMGSRGSSAYSRLRDSVSAVL 64

QY 102 CTLIVALGPIQFGFTSGFSSPTQDAMVRDLNLSISEFSAFGSLSNVGMVGAIASGQMAE 161
Db 65 CTLIVALGPIQFGFTCGFSSPTQDAISDLGLTLSEFSLFGLSNVGMVGAIASGQIAE 124

QY 162 YIGRKGSMLMAAIPNIIIGWLAIISFAKDASFLYMGRLLEGFGVGIISYTPVYIAEISPN 221
Db 125 YIGRKGSMLMAAIPNIIIGWLAIISFAKDS-----XVPVYIAEIAPO 165

QY 222 MRGALGSVNQLSVTFGIFLAYLLGMFIPWRLLAVIGALPCTMLIPGLFFIPESPRWLAKM 281
Db 166 MRGALGSVNQLSVTIGILLAYLLGMFVPWRILSVLILPCSILIPGLFFIPESPRWLAKM 225

QY 282 NLTEDCETSLQVIRGFETDITTEVNDIKRAVASSSKRTTISFQELNQKKYRTPLLLGIGL 341
Db 226 GKMEDFESSLQVIRGFETDIAVEVNEIK-----IGIGL 258

QY 342 LVLQNLSGINGVLFYASSIFKAAGVTNSDLATCSLGAIQVLATGVTWLLDRAGRILLI 401
Db 259 LVLQQLSGVNGILFYAASIFKAAGLTNSNLATFGLGVVQ----- 297

QY 402 ISTSGMTCLLAVSVVFFFLKDNISQDSNSVYILTMISLVGIVSFVITFSFGMAIPWLM 461
Db 298 ISTTGMTITLVVSVVSFFVKDNITNGSHLYSVMSLSLVGLV----- 339

QY 462 SEILPVSIKSLGGSIAATLANWLTSFAITMTNMLTWSVGGTFLSYMVVSFTIVFVVLW 521
Db 340 --ILPVDIKSLAGSVATLANWLTAWLITMTASMLWSNGGTFATIAAFCAGTLVFCVCLW 397

QY 522 VPETKG 527
Db 398 VPETKG 403

RESULT 9
US-10-425-114-65425
; Sequence 65425, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 65425
; LENGTH: 318
; TYPE: PRT
```

```
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4764-010-D11_FLI.pep
US-10-425-114-65425

Query Match      43.9%; Score 1219; DB 15; Length 318;
Best Local Similarity 75.9%; Pred. No. 2.1e-97;
Matches 246; Conservative 24; Mismatches 38; Indels 16; Gaps 5;

QY 9 PAGKKKKKXNQLRREAVPGRPASBELRTRVMGGGSRGGAGAGE-ESGSDHDGVLRRL 67
Db 6 PTRRKRKNPNL-----AASWAARARGAMARSRTSGGGEYESGSDHDGALQKPLL 54

QY 68 -NTGSWYR--MSSRQSSF-APGTSSMAVLRESHVSAFLCTLIVALGPIQFGFTSGFSSPT 123
Db 55 PMSGWYRMGMGRQSSLNAAAGTSSMAVLRESHVSAFLCTLIVALGPIQFGFTGGYSSPT 114

QY 124 QDAMVRDLNLSISEFSAFGSLSNVGMVGAIASGQMAEYIGRKGSMLMAAIPNIIIGWLAI 183
Db 115 QASIIIRDNLNLSISEFSVFGSLSNVGMVGAIASGQMAEYVGRKGSMLMAAIPNIIIGWLAI 174

QY 184 SFAKDASFLYMGRLLEGFGVGIISYTPVYIAEISPNMRGALGSVNQLSVTFGIFLAYL 243
Db 175 SFAKDSSFLYMGRLLEGFGVGIISYVVPVYIAEISPNMRGALGSVNQLSVTLGIMFAYL 234

QY 244 LGMFIPWRLLAVIGALPCTMLIPGLFFIPESPRWLAKMNLTEDCETSLQVIRGFETDITT 303
Db 235 LGLFVPWRLLAVIGLIPCLIVLIPGLFFIPESPRWLAKNMMDDCETSLQVIRGFADADITA 294

QY 304 EVNDIKRAVASSSKRTTISFQELN 327
Db 295 ELNDIKRAVMSANKRATIRFQELN 318

RESULT 10
US-10-424-599-275268
; Sequence 275268, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 275268
; LENGTH: 361
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(361)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_90588C.1.pep
US-10-424-599-275268

Query Match      43.1%; Score 1199; DB 15; Length 361;
Best Local Similarity 65.1%; Pred. No. 1.4e-95;
Matches 226; Conservative 63; Mismatches 52; Indels 6; Gaps 1;

QY 187 KDASFLYMGRLLEGFGVGIISYTPVYIAEISPNMRGALGSVNQLSVTFGIFLAYLLGM 246
Db 4 QDSSFLYMGSLLEGFGVGIISYXVPVYIAEIAPNLRGGLGSVNQLSITIGIMLAYLLGL 63

QY 247 FIPWRLLAVIGALPCTMLIPGLFFIPESPRWLAKMNLTEDCETSLQVIRGFETDITTEVN 306
Db 64 FVNWRVLAAILGILPCTVLIPGLFFIPESPRWLAKMGTDEFETSLQVIRGFDTDISVEVY 123
```

```
QY 307 DIKRAVASSKRTTISFQELNQKRYTPLL-----LGIGLLVLQNLSGINGVLFYASSI 360
Db 124 EIKRSVASTGKRATIRFADLKRKYWFPLMVXYDVTVGIGLLVLQQLSGINGVLFYSTTI 183
QY 361 FKAAGVTNSDLATCSLGAIOVLATGVTTWLLDRAGRRILLIISTSGMTLCLLAVSVVFFL 420
Db 184 FANAGISSEAAATVGLGAVQVIATGISTWLVDKSGRRLLMISSVMTVSLLIVSIAYFL 243
QY 421 KDNISQDNSYYILTMISLVGIVSFVITFSFGMGAIPWLMNSEILPVSIKSLGGSIAITLA 480
Db 244 EGVVSEDSHLFSLMGLIVSVVGLVVMVIGFSLGLGPIPLIMSEILPVNIKGLAGSIATMG 303
QY 481 NWLTSFAITMTNMLTWSVGGTFLSYMVVSFAITIVFVVLWVPETKG 527
Db 304 NWLISWVITMTANLLLNWNSGGTFTIYTVVAAFTIAFTALWVPETKG 350

RESULT 11
US-10-424-599-273908
; Sequence 273908, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 273908
; LENGTH: 338
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_8935C.1.pep
US-10-424-599-273908

Query Match 42.6%; Score 1185; DB 15; Length 338;
Best Local Similarity 64.3%; Pred. No. 2.1e-94;
Matches 223; Conservative 61; Mismatches 51; Indels 12; Gaps 2;

QY 201 FGVGIIISYTPVYIAEISPPNMGALGSVNQLSVTFGIFLAYLLGMFIPWRLAVIGALP 260
Db 1 FGVGIIISYTPVYIAEISPPNMGALGSVNQLSVTFGIFLAYLLGMFIPWRLAVIGALP 60
QY 261 CTMLIPGLFFIPESPRWLAKMNLTEDCETSLQVLRGPETDITTEVNDIKRAVASSKRTT 320
Db 61 CTILIPGLFFIPESPRWLAKMNGTEEFETSLQVLRGPETDISVEVEIKRAVASTNRRIT 120
QY 321 ISFQELNQKRYTPLLIGIGLLVLQNLSGINGVLFYASSIFKAAGVTNSDLATCSLGAIQ 380
Db 121 VRFADLKORRYWLPLMIGIGLLVLQQLSGINGVLFYSSITFRNAGISSSDAATFGVGAVQ 180
QY 381 VLATGVTTWLLDRAGRRILLIISTSGMTLCLLAVSVVFFLKDNIQSNSYYILTMISLV 440
Db 181 VLATSLLWLADKSGRRLLIVSATGMSFSLVVAITFYIKASISSETSSLYGILSTLSLV 240
QY 441 GIVSFVITFSFGMGAIPWLMNSEILPVSIKSLGGSIAITLANWLTSAITMTNMLTWSV 500
Db 241 GVVAMVIAFSLGMGAMPWIIMSEILPINIKGLAGSVATLANWLFSLVTLTANMLLDWSS 300
QY 501 GGTFLSYMVVSFAITIVFVVLWVPETKGXNSRGDTIFVLSLIQRLQW 547
Db 301 GGTFTIYAVVCALTWVFVTWVPETK-----GKTI-----EEIQW 335

RESULT 12
US-10-425-114-49786
; Sequence 49786, Application US/10425114
; Publication No. US20040034888A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 49786
; LENGTH: 219
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700464574_FLI.pep
US-10-425-114-49786

Query Match 37.1%; Score 1032; DB 15; Length 219;
Best Local Similarity 100.0%; Pred. No. 2.3e-81;
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 320 TISFQELNQKRYTPLLIGIGLLVLQNLSGINGVLFYASSIFKAAGVTNSDLATCSLGA 379
Db 1 TISFQELNQKRYTPLLIGIGLLVLQNLSGINGVLFYASSIFKAAGVTNSDLATCSLGA 60
QY 380 QVLATGVTTWLLDRAGRRILLIISTSGMTLCLLAVSVVFFLKDNIQSNSYYILTMISL 439
Db 61 QVLATGVTTWLLDRAGRRILLIISTSGMTLCLLAVSVVFFLKDNIQSNSYYILTMISL 120
QY 440 VGIVSFVITFSFGMGAIPWLMNSEILPVSIKSLGGSIAITLANWLTSAITMTNMLTWS 499
Db 121 VGIVSFVITFSFGMGAIPWLMNSEILPVSIKSLGGSIAITLANWLTSAITMTNMLTWS 180
QY 500 VGGTFLSYMVVSFAITIVFVVLWVPETKG 527
Db 181 VGGTFLSYMVVSFAITIVFVVLWVPETKG 208

RESULT 13
US-10-437-963-166774
; Sequence 166774, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 166774
; LENGTH: 502
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_6544C.1.pep
US-10-437-963-166774

Query Match 32.7%; Score 909.5; DB 16; Length 502;
Best Local Similarity 37.0%; Pred. No. 3.2e-70;
Matches 194; Conservative 101; Mismatches 158; Indels 71; Gaps 8;
```

QY 29 RPASELRTFWMGGSNRGGAGAGEESGDHGDGVLRRPLL-----NTGSW 72
Db 5 RAAEDVEGAVVAGAAGGGGGG-----GEVTAPLLLRQHKQGRGDEEKIQNDAGR 56
QY 73 YRMSSRQSSPAPGTSSMAVLRSHVSAFLCTLIIVALGP IQFGFTSGFSSPTQDAMVRDLN 132
Db 57 LRGSARR-----KRRRV-----GYSAPTQSGIVDEVG 83
QY 133 LSISEFSAFGSLSNVGMVGAIASGQMAEYIGRK-----GSLMIAAIPNIIGWLAI 183
Db 84 LSISQFALFGSVLTIGAMIGAVTSGRLADFLGRKMVCHIFRLSQTMRIASATICIFGWL 143
QY 184 SFAKDASFLYMGRLLEGFGVGIIISYTPVYIAEISPNMRGALGSVNQLSVTFGIFLAYL 243
Db 144 HLAGKVIMLYFGRILLGFSTGILSYVVPVFI AEIAPKNLRGGLATSNQLLICSGSSATYI 203
QY 244 LGMFIPWRLLAVIGALPCTMLIPGLFFIPESPRWLAKMNLTEDCETSLQVLRGFETDIT 303
Db 204 IGA LVARNLVLVGIVPCVLLLTGLLFPESPRWLANVGREKHFHASLQMLRGEDADVSE 263
QY 304 EVNDIKRAVASSSKRTTISFQELNQKKYRTPLLLGIGLLVLQNLSGINGVLFYASSIFKA 363
Db 264 EAVEIKEYIESLHRFPKARVQDLFLRKNIYAVTVGVGLMIFQQLGGINGVGFYASSIFTS 323
QY 364 AGVTNSDLATCSLGAIQVLATGVTWTLDRAGRRILLIISTSGMTLCLLAVSVVFFLKD 423
Db 324 AGPSGK-LGTILIGIIQIPITLFGAILMDKSGRRVLLMVSASGTLGCFLTGISFYLK-- 380
QY 424 ISQDSNSYVILTMISLVGIVSFVITFSFGMGAIPWLMMSSEILPVSIKSLGGSIA TLANWL 483
Db 381 -AQGLFSEWV-PELALTGILVYIGAYSIGMGVPVWVMSEIFSIDMKAIGGSLVTLVSWL 438
QY 484 TSFAITMTNMLTWSVGGTFSLSYMVSAPTIVFVVLWVPETKG 527
Db 439 GSFAISYFSFLMDWSSAGTFFMFSAASLITILFVVMVVPETKG 482

RESULT 14

US-10-310-154-725
; Sequence 725, Application US/10310154
; Publication No. US20030233670A1
; GENERAL INFORMATION:

; APPLICANT: Edgerton, Michael D
; APPLICANT: Chomet, Paul S.
; APPLICANT: Adams, Thomas H
; APPLICANT: Ruff, Thomas G.
; APPLICANT: Agarwal, Ameeta K.
; APPLICANT: Ahrens, Jeffrey E.
; APPLICANT: Ball, James A.
; APPLICANT: Banu, G.
; APPLICANT: Bell, Erin
; APPLICANT: Boddupalli, Raghava
; APPLICANT: Deikman, Jill
; APPLICANT: Deng, Molian
; APPLICANT: Dong, Jinzhao
; APPLICANT: Duff, Stephen M.
; APPLICANT: Galligan, Meghan M.
; APPLICANT: Hinchey, Brenda S.
; APPLICANT: Huang, Shihshieh
; APPLICANT: Johnson, G. Richard
; APPLICANT: Jung, Vincent
; APPLICANT: Kretzmer, Keith A
; APPLICANT: Laccetti, Lucille B.
; APPLICANT: Lai, Chao-Qiang
; APPLICANT: Lee, Gary
; APPLICANT: Lin, Jie-Yi
; APPLICANT: Liu, Jingdong
; APPLICANT: Lu, Bin
; APPLICANT: Luethy, Michael M.
; APPLICANT: Lund, Adrian
; APPLICANT: Madson, Linda L.
; APPLICANT: Malloy, Kathleen A.
; APPLICANT: McKiel, Christine L.

; APPLICANT: Miller, Philip W.
; APPLICANT: Padmavathi, Manchikanti
; APPLICANT: Parnell, Laurence D.
; APPLICANT: Start, William G.
; APPLICANT: Tennesen, Dan
; APPLICANT: Vidya, K.R.
; APPLICANT: Wang, Haiyun
; APPLICANT: Xin, Zhanguo
; APPLICANT: Xu, Nanfei
; APPLICANT: Yang, Chunzhi
; APPLICANT: Zeng, Xiaoping
; APPLICANT: Zhang, Qiang
; APPLICANT: Zhao, Yajuan
; APPLICANT: Zhou, Li
; TITLE OF INVENTION: Gene Sequences and Uses Thereof in Plants
; FILE REFERENCE: 38-15(52796)B
; CURRENT APPLICATION NUMBER: US/10/310,154
; CURRENT FILING DATE: 2002-12-04
; PRIOR APPLICATION NUMBER: 60/337,358
; PRIOR FILING DATE: 2001-12-04
; NUMBER OF SEQ ID NOS: 736
; SEQ ID NO 725
; LENGTH: 481
; TYPE: PR1
; ORGANISM: Glycine max
; US-10-310-154-725

Query Match 32.1%; Score 891; DB 14; Length 481;

Best Local Similarity 37.2%; Pred. No. 1.2e-68;

Matches 188; Conservative 97; Mismatches 171; Indels 50; Gaps 4;

QY 48 AGAGEESGDHGDGVLRRPLLNTGWSYRMSSRQSSPAPGTSSMAVLRSHVSAFLCTLI 107
Db 25 AAACKENGSD-----KSVKNGSIGMVL-----LSTLVAV 53
QY 108 LGPIQFGFTSGFSSPTQDAMVRDLNLSISEPSAFGSLSNVGMVGAIASGQMAEYIGRK 167
Db 54 CGSFTFGTCVGSAPTQAAIRADNLNLSAEFSNMGSLVTIGAMLGAITSGRITDFIGRK 113
QY 168 SLMIAAIPNIIGWLAI SFAKDASFLYMGRLLEGFGVGIIISYTPVYIAEISPNMRGALG 227
Db 114 AMRISTGFCITGWIAVFFSKGSYSLDFGRFTGYGIGVISYVVPVYIAEIA PKNLRGGLA 173
QY 228 SVNQLSVTFGIFLAYLLGMFIPWRLLAVIGALPCTMLIPGLFFIPESPRWLAKNLTEDC 287
Db 174 TTNQLLIVTGGSVFLLGSVINWRRELALAGLVPICLLVGLCFIPESPRWLAKVGREKEF 233
QY 288 ETSLQVLRGFETDITTEVNDIKRAVASSSKRTTISFQELNQKKYRTPLLLGIGLLVLQNL 347
Db 234 QLALSRLRGKDADISDEAAEILDYIETLQSLPKTKLLDLFQSKYVHSVIVGVGLMACQQS 293
QY 348 SGINGVLFYASSIFKAAGVTNSDLATCSLGAIQVLATGVTWTLDRAGRRILLIISTSGM 407
Db 294 VGINGIGFYTAEIFVAAGLSSGKAGTIAYACIQIPFTLLGAILMDKSGRRPLVMVSAAGT 353
QY 408 TLCLLAVSVVFFLKDNTSQDSNSYYILTMISLVGIVSFVITFSFGMGAIPWLMMSSEILPV 467
Db 354 FLCGFVAFAFFLKD----QSLPEWVPILAFAGVLIIYIAAFSIGLSVVPWIMSEIFPI 409
QY 468 SIXSLGGSIA TLANWLTSFAITMTNMLTWSVGGTFSLSYMVSAPTIVFVVLWVPETKG 527
Db 410 HLKGTAGSLVVLVAVLGA WVVSYTFNFLMSWSSPGTLFLYAGCSLLTILFVAKLVPETKG 469
QY 528 XNSRGDTIFVSLSIQRLQWLPECLS 553
Db 470 -----KTLEEIQACIS 480

RESULT 15

US-10-424-599-275267
; Sequence 275267, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 275267
; LENGTH: 216
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(216)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_90587C.1.pap
US-10-424-599-275267

Query Match 27.6%; Score 768; DB 15; Length 216;
Best Local Similarity 75.0%; Pred. No. 2.1e-58;
Matches 156; Conservative 17; Mismatches 31; Indels 4; Gaps 2;

QY	52	EESGSDH	DGVLRRLN	TGSWYRMS	RQSSFAP	GTSSMA	VLRSHV	SAFLCT	LIV	ALGPI	111
Db	7	EESGDAR	--LQK	PFLHTG	SWYKMS	RQSSIM	--GS	THVIRD	GAVSV	LCVLI	ALGPI 62
QY	112	QFGFTG	FSSPTQ	DAMVRD	NLSISE	FSAPG	SLSNV	GMVGA	IASG	QMAEY	IGRKGS
Db	63	QFGFTG	GYSSPT	QGAIVR	DLNLSI	SEFSF	FGSL	SNVGA	MVGA	IASG	QIAEY
QY	172	AAIPNI	IGWLA	ISFAK	DA	SFLY	MGRLL	EGFV	GIISY	TVPV	YIAE
Db	123	AAIPNI	IGWLA	ISFAK	DS	FLYX	GRLL	EGFV	FC	SVXV	PVYIAE
QY	232	LSVTFG	IFLAYL	LG	MFIP	WRLL	AVIGAL	259			
Db	183	LSVTIG	IMLAYL	LLGL	FVNR	VRVLA	ILGKL	210			

Search completed: October 13, 2004, 11:54:39
Job time : 210.265 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 13, 2004, 11:39:29 ; Search time 6383 Seconds
(without alignments)
4097.007 Million cell updates/sec

Title: US-10-051-909-36

Perfect score: 2779

Sequence: 1 PSSSSSFRPAGKKKKKNQG.....TIFVSLSIQRQLWLPECLS 553

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool_h/US10051909/runat_13102004_123336_19906/app_query.fasta_1.1678
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10051909 @CGN_1_1_9091 @runat_13102004_123336_19906 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_in:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2133	76.8	1870	8 AK100792	AK100792 Oryza sat
2	2017.5	72.6	2266	8 AK065497	AK065497 Oryza sat
3	1821.5	65.5	1948	8 AK068594	AK068594 Oryza sat
4	1757.5	63.2	2020	8 BVU43629	U43629 Beta vulgar

5	1746.5	62.8	1557	8 AY093274	AY093274 Arabidops
6	1746.5	62.8	1866	8 AY059848	AY059848 Arabidops
7	1738.5	62.6	1695	8 AF412060	AF412060 Arabidops
8	1734.5	62.4	1464	8 AY124845	AY124845 Arabidops
9	1726.5	62.1	1914	8 AF462803	AF462803 Arabidops
10	1710	61.5	1377	6 AX653119	AX653119 Sequence
11	1693	60.9	1825	8 BT014234	BT014234 Lycopersi
12	1577	56.7	1428	6 AX653120	AX653120 Sequence
13	1453.5	52.3	204649	8 AC120986	AC120986 Oryza sat
14	1137.5	40.9	120977	8 AC025808	AC025808 Genomic s
15	1117.5	40.2	68041	8 AC025814	AC025814 Arabidops
16	992.5	35.7	99587	8 AC007858	AC007858 Oryza sat
17	982	35.3	156643	8 AC120988	AC120988 Oryza sat
18	965.5	34.7	1699	8 LES278765	AJ278765 Lycopersi
19	954.5	34.3	2008	8 AK106658	AK106658 Oryza sat
20	951.5	34.2	1741	8 AY084684	AY084684 Arabidops
21	944.5	34.0	1778	8 AY064144	AY064144 Arabidops
22	943.5	34.0	1435	8 BT000053	BT000053 Arabidops
23	943.5	34.0	1686	8 AY120715	AY120715 Arabidops
24	939	33.8	1449	8 BT000608	BT000608 Arabidops
25	935.5	33.7	1747	8 AY087180	AY087180 Arabidops
26	932	33.5	1806	8 BT008661	BT008661 Arabidops
27	932	33.5	1847	8 ATH249967	AJ249967 Arabidops
28	920.5	33.1	1473	6 AX507636	AX507636 Sequence
29	920.5	33.1	1678	8 D89051	D89051 Arabidopsis
30	899	32.3	1359	6 AX506932	AX506932 Sequence
31	898.5	32.3	1420	8 AY091216	AY091216 Arabidops
32	898.5	32.3	1660	8 AY063856	AY063856 Arabidops
33	891.5	32.1	1755	8 AK064996	AK064996 Oryza sat
34	878.5	31.6	1729	8 AF367260	AF367260 Arabidops
35	877.5	31.6	1407	6 AX506596	AX506596 Sequence
36	877.5	31.6	1407	6 AX651294	AX651294 Sequence
37	877.5	31.6	1413	8 AY133547	AY133547 Arabidops
38	870.5	31.3	1606	8 AY026255	AY026255 Arabidops
39	866	31.2	1578	8 ATH249968	AJ249968 Arabidops
40	858.5	30.9	1757	8 AY048207	AY048207 Arabidops
41	858.5	30.9	1757	8 AY123991	AY123991 Arabidops
42	853.5	30.7	1787	8 AY026254	AY026254 Arabidops
43	845	30.4	1668	8 AY058056	AY058056 Arabidops
44	843.5	30.4	1377	8 BT002699	BT002699 Arabidops
45	837.5	30.1	1209	6 AX653566	AX653566 Sequence

ALIGNMENTS

RESULT 1	AK100792	Oryza sativa (japonica cultivar-group)	1870 bp	mrna	linear	PLN 24-JUL-2003
AK100792	Oryza sativa (japonica cultivar-group)	cdna clone:J023121D20, full				
LOCUS	AK100792	insert sequence.				
DEFINITION	AK100792	GI:32986001				
ACCESSION	AK100792	FLI_CDNA; CAP trapper.				
VERSION	AK100792.1	Oryza sativa (japonica cultivar-group)				
KEYWORDS	FLI_CDNA; CAP trapper.	Oryza sativa (japonica cultivar-group)				
SOURCE	AK100792	Oryza sativa (japonica cultivar-group)				
ORGANISM	AK100792	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.				
REFERENCE	1	The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-length cDNA Project Team; Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group; Otomo,Y., Murakami,K., Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y., Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M., Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J., Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J., Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN; Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M.,				
AUTHORS						

QY	314	SerSerSerLysArgThrThrIleSerPheGlnGlnLeuAsnGlnLysLysTyrArgThr	333
Db	999	TCAGCAAAACAAAGGACACGATCCGTTTCAAGAATTAAACCAGAGAAATACCGCAC	1058
QY	334	ProLeuLeuGlyIleGlyLeuValLeuGlnAsnLeuSerGlyIleAsnGlyVal	353
Db	1059	CCCCTAATACTAGGAATTGGCCTACTGTACTGCAACAGCTAAGTGAATCAATGGAATA	1118
QY	354	LeuPheTyrAlaSerSerIlePheLysAlaAlaGlyValThrAsnSerAspLeuAlaThr	373
Db	1119	TTGTTTTATGCAGGTAGCATCTTCAAAGCAGCAGGCTCTCAAAACAGTGACTGGGTACA	1178
QY	374	CysSerLeuGlyAlaIleGlnValLeuAlaThrGlyValThrThrTlpLeuLeuAspArg	393
Db	1179	TGTGCACCTGGTGCTATCCAGGTTCTTGCTACAGGAGTTACAACCTGGTTATTAGACAGA	1238
QY	394	AlaGlyArgArgIleLeuLeuIleIleSerThrSerGlyMetThrLeuCysLeuLeuAla	413
Db	1239	GCTGGCCGCGGATCCTCTTATCATCTCTTCTGCTGGGATGACTCTAAGCCTCCTTGCA	1298
QY	414	ValSerValValPhePheLeuLysAspAsnIleSerGlnAspSerAsnSerTyrTyrIle	433
Db	1299	GTTGCTGTTGTTATTTTCTCAAGGATAGCATTTCAACAAGATTCTCACATGTACTACACC	1358
QY	434	LeuThrMetIleSerLeuValGlyIleValSerPheValIleThrPheSerPheGlyMet	453
Db	1359	TTAAGTATGATCTCCTTGGTCTCTGTGGCTTTTGTATATCGCTTCTCCTTCGGTATG	1418
QY	454	GlyAlaIleProTrpLeuMetMetSerGluIleLeuProValSerIleLysSerLeuGly	473
Db	1419	GGTGCCATTCATGGATCATATAATGTCAGAGATCCTCCCGGTTAGTATCAAGAGTCTCGCA	1478
QY	474	GlySerIleAlaThrLeuAlaAsnTrpLeuThrSerPheAlaIleThrMetThrThrAsn	493
Db	1479	GGAAGCTTTGGCAGCTCGCCAACTGGCTTACATCCTTTGGATAACAATGACAGCAAC	1538
QY	494	LeuMetLeuThrTrpSerValGlyGlyThrPheLeuSerTyrMetValValSerAlaPhe	513
Db	1539	TTGATGCTTAGCTGGAGTGCTGGAGGACCTTGTGCTCTACATGCTGAGTGCTTTC	1598
QY	514	ThrIleValPheValValLeuTrpValProGluThrLysGly***AsnSerArgGlyAsp	533
Db	1599	ACCCTCGTGTTCGTCATCCTTTGGTGCCAGACAAAGGAAG-AACTCTCGAAGAGAT	1657
QY	534	ThrIlePheValSerLeuSerIle	541
Db	1658	ACAATGGTCTTCGCTGAGCCTT	1681
RESULT 2			
AK065497			
LOCUS			
DEFINITION	Oryza sativa (japonica cultivar-group) cDNA clone:J013023P09, full insert sequence.		
ACCESSION	AK065497		
VERSION	AK065497.1 GI:32975515		
KEYWORDS	FLI CDNA; CAP trapper.		
SOURCE	Oryza sativa (japonica cultivar-group)		
ORGANISM	Oryza sativa (japonica cultivar-group)		
REFERENCE	1 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.		
AUTHORS	The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-Length cDNA Project Team; Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T. and Yamamoto,M.		

TITLE	japonica rice		
JOURNAL	Science 301 (5631), 376-379 (2003)		
MEDLINE	22752273		
PUBMED	12869764		
REFERENCE	2 (bases 1 to 2266)		
AUTHORS	Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K., Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K., Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M., Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M., Kodama,T., Kojima,K., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Kurosaki,T., Kusumegi,T., Li,C., Lu,M., Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A., Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakamura,M., Namiki,T., Narikawa,R., Niikura,J., Nishi,K., Nomura,K., Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H., Osato,N., Ota,Y., Otomo,Y., Ryu,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shinagawa,T., Sugano,S., Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S., Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,W., Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S. and Yoshimura,A.		
TITLE	Direct Submission		
JOURNAL	Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)		
COMMENT	This clone is one of the 28K full-length cDNA clones from japonica rice. URL : http://cdna01.dna.affrc.go.jp/cdna/ NIAS Rice Full-Length cDNA Project Team: Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T. and Yamamoto,M. FAIS Genome Sequencing & Analysis Group: Otomo,Y., Iida,Y., Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M., Kodama,T., Kurosaki,T., Kusumegi,T., Lu,M., Masuda,H., Miura,J., Mizuno,K., Narikawa,R., Niikura,J., Oka,M., Ryu,R., Sugano,S., Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S., Yoshimura,A., Matsubara,K. and Murakami,K. Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Iida,J., Imamura,K., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,J., Kishikawa-Hirozane,T., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Osato,N., Ota,Y., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Waki,K., Yasunishi,A. and Hayashizaki,Y. Location/Qualifiers 1. .2266 /organism="Oryza sativa (japonica cultivar-group)" /mol type="mRNA" /cultivar="Nipponbare"		

Db 229 ---GGAGATCTGGCGAAACCAATTTTGGCATACAGGAAGTTGGTATAGAAATGGGTCCAGA 285

QY 79 GlnSerSerPheAlaProGlyThrSerSerMetAlaValLeuArgGluSerHisValSer 98

Db 286 CAATCAAGTTTG-----ATGGGTTCTCTCTCAAGTTATTTCGAGAAAGCTCCATCTCT 336

QY 99 AlaPheLeuCysThrLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGly 118

Db 337 GTCCTTGCTTGTTCTTATTGTTGCTCTTGCTCTTATTCAGTTTGGTTTACTGCTGGT 396

QY 119 PheSerSerProThrGlnAspAlaMetValArgAspLeuAsnLeuSerIleSerGluPhe 138

Db 397 TATTTCATCTCCAATCAATCAGCAATTAACAATGAACCTTGGTCTTCTTGTGCAGAGTAC 456

QY 139 SerAlaPheGlySerLeuSerAsnValGlyGlyMetValGlyAlaIleAlaSerGlyGln 158

Db 457 TCATGGTTTGGTCTCTATCCAAATGTGGGTGCCATGGTTGGGCAATTGCAAGTGGTCAA 516

QY 159 MetAlaGluTyrIleGlyArgLysGlySerLeuMetIleAlaAlaIleProAsnIleIle 178

Db 517 ATATCTGAGTATATTGGAAGAAAGGGTCTTTGATGATTTGCTGCTATCCCCAACATAATT 576

QY 179 GlyTyrLeuAlaIleSerPheAlaLysAspAlaSerPheLeuTyrMetGlyArgLeuLeu 198

Db 577 GGATGGCTTGCTATATCATTTGCCAAAGATTCTTCATTTCTGTATATGGGAAGGATGTG 636

QY 199 GluGlyPheGlyValGlyIleIleSerTyrThrValProValTyrIleAlaGluIleSer 218

Db 637 GAAGGTTTTGGCGTGGGAATAATATCTTACACGGTTCCTGTGTACATTCTGAAATAGCA 696

QY 219 ProGlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIle 238

Db 697 CCACAGAAATTTGAGGGGAGCTTTGGGTTTCAGTAAACCAGCTATCCGTTACAATCGGTATA 756

QY 239 PheLeuAlaTyrLeuLeuGlyMetPheIleProTyrArgLeuLeuAlaValIleGlyAla 258

Db 757 ATGCTATCGTATATGCTGGGGCTTTTGTCCCGTGGAGGATACCTTGACGTTTATAGGAATT 816

QY 259 LeuProCysThrMetLeuIleProGlyLeuPhePheIleProGluSerProArgTyrLeu 278

Db 817 TTACCATGCACATATATGATACCTGGCTTATCTTCATACCGGAATCTCCTAGGTGGTTG 876

QY 279 AlaLysMetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPheGlu 298

Db 877 GCGAAAATGGGTATGATGGAAGAAATTTGAGGTTTCTTGCAGATTCTCCGGGGCTTTGAC 936

QY 299 ThrAspIleThrThrGluValAsnAspIleLysArgAlaValAlaSerSerLysArg 318

Db 937 ACTGATATATCTCTGAAGTTAATGAATCAAGATCTGTTGCTTCAAGTTCTCCGGGGCTTTGAC 996

QY 319 ThrThrIleSerPheGlnGluLeuAsnGlnLysLysTyrArgThrProLeuLeuLeuGly 338

Db 997 ACCACCATAGAATTTGCAGAACTACGACAAAGAGATATTGGTTACCCTTGATGATAGG 1056

QY 339 IleGlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSer 358

Db 1057 AATGGTTTACTTATTTTGCAACAACCTCAGTGGAAATCAATGGTGTCTCTTATCTAGT 1116

QY 359 SerIlePheLysAlaAlaGlyValThrAsnSerAspLeuAlaThrCysSerLeuGlyAla 378

Db 1117 ACCATCTTTAAAGAGGCTGGGTCACATCAAGCAACGCTGCAACATTTGGCCTTGGTGCT 1176

QY 379 IleGlnValLeuAlaThrGlyValThrThrTyrLeuLeuAspArgAlaGlyArgArgIle 398

Db 1177 GTTCAGGTTATCGCAACTGTTGTGACTACCTGGTTGGTGGATAAATCAGACGCAGACTT 1236

QY 399 LeuLeuIleIleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSerValValPhe 418

Db 1237 CTTCTAATTGCTCTTCTTCTGGGATGACATTGAGTCTTCTAGTGGTTCGCAATGTCCTTC 1296

QY 419 PheLeuLysAspAsnIleSerGlnAspSerAsnSerTyrTyrIleLeuThrMetIleSer 438

Db 1297 TTTTAAAGGAAATGGTATCAGACGAATCTACATGTCAGTGTATTCTCCATCCTTTCA 1356

QY 439 LeuValGlyIleValSerPheValIleThrPheSerPheGlyMetGlyAlaIleProTyr 458

Db 1357 GTTGTGGAGTTGTGGCTATGGTGGTTACTTCTCACTAGGAATTGGAGCAATTCCTGG 1416

QY 459 LeuMetMetSerGluIleLeuProValSerIleLysSerLeuGlyGlySerIleAlaThr 478

Db 1417 ATTATAATGTCAAGATTCTTCCAATTAACATCAAGGGCTGGCTGGAAGTATTGCTACA 1476

QY 479 LeuAlaAsnTyrLeuThrSerPheAlaIleThrMetThrThrAsnLeuMetLeuThrTyr 498

Db 1477 CTTGCCAACTGGTTTGAGCATGGATTGTGACAAATGACTGCAACATAATGTTGTCATGG 1536

QY 499 SerValGlyGlyThrPheLeuSerTyrMetValValSerAlaPheThrIleValPheVal 518

Db 1537 AATAGTGGAGGTACTTTCTCTATTACATGGTTGTTGTGCTTTCACGTGGCCTTTGTG 1596

QY 519 ValLeuTyrValProGluThrLysGly***AsnSerArgGlyAspThrIlePheValSer 538

Db 1597 GTCATTTGGTCCCGAGACAAAGGGAG-AACATTGGAGGAGATTTCAGTG---GTCCTT 1652

QY 539 LeuSerIleGlnArgGlnLeuGlnTyrLeuProGluCys 551

Db 1653 CAGAAGATAAAGAAATTTCTGATAACTTGTCCATGGTGT 1691

RESULT 5

AY093274

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AY093274 1557 bp mRNA linear PLN 21-APR-2002

Arabidopsis thaliana similar to integral membrane protein (Atig19450) mRNA, complete cds.

AY093274 GI:20259851

FLJ CDNA.

Arabidopsis thaliana (thale cress)

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

1 (bases 1 to 1557)

Southwick,A., Karlin-Neumann,G., Nguyen,M., Lam,B., Miranda,M., Palm,C.J., Bowser,L., Jones,T., Banh,J., Carninci,P., Chen,H., Cheuk,R., Chung,M.K., Hayashizaki,Y., Ishida,J., Kamiya,A., Kawai,J., Kim,C., Lin,J., Liu,S.X., Narusaka,M., Pham,P.K., Sakano,H., Sakurai,T., Satou,M., Seki,M., Shinn,P., Yamada,K., Shinozaki,K., Ecker,J., Theologis,A. and Davis,R.W.

Direct Submission

Submitted (26-MAR-2002) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA

e-mail for correspondence: arab@sequence.stanford.edu

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-length cDNA'): Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PCEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Southwick,A., Nguyen,M., Karlin-Neumann,G., Lam,B., Miranda,M., Palm,C.J., Bowser,L., Jones,T., Banh,J., Chen,H., Cheuk,R., Chung,M.K., Kim,C., Lin,J., Liu,S.X., Pham,P.K., Sakano,H., Shinn,P., Yamada,K., Ecker,J., Theologis,A. and Davis,R.W.

Southwick,A., (SSP/Stanford) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Davis,R.W. (SSP/Stanford) contributed equally to this work as PIs.

Location/Qualifiers

1.1557

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/chromosome="1"

FEATURES

source

gene
1. .1557
/clone="U16352"
/ecotype="Columbia"
/note="This clone is in pENTR/SD-dTopo This is a cloned PCR product using RIKEN clone RAFLO7-13-K02 (AY059848) as a template"
CDS
1. .1467
/gene="At1g19450"
/note="synonym: F18014.22"
1. .1467
/gene="At1g19450"
/codon_start=1
/product="similar to integral membrane protein"
/protein_id="AAM13273.1"
/db_xref="GI:20259852"
/translation="MSFRDNDTEEGRNDRRLRPFLHTGWSYRMGSRQSSMLESSQVIRD
SSISVLACVLIVAGLPIQFGFTCGYSSTQAAITKDLGLTVSEYSVFGSLSNVGAMVG
ATASGQIAEYVGRKGSMLIAIPNIIGWLSISPAKDTSPFLYMGRLLEGGVGIISYTV
PVYIAEIAPOQTMRGALGSVNQLSVTIGIMLAYLGLFVWPRIILAVLGLPCTLLIPGL
FFIPESPRWLAKMGLTDDPETSLOVLRGFETDITVEVNEIKRSVASSKRSARVFDL
KRRYYFPLMVGIGLLAQQLGGINGVLFYSSITPESAGVTSNVATFGVGVQVWAT
GIATWLVDKAGRRLLMISIGMTISLVIIVAVAFYLKEFVSPDSNMVNIILSMVSVVG
VAMVISCSLGMGPIPLWIMSEILPVNIKGLAGSIATLLNWFVSWLVTMTANMLLAWSS
GGTFTLYALVCGFTTVFVSLWVPETKGTLEIQAIFR"

ORIGIN

Alignment Scores:
Pred. No.: 1.98e-121 Length: 1557
Score: 1746.50 Matches: 338
Percent Similarity: 83.40% Conservative: 69
Best Local Similarity: 69.26% Mismatches: 74
Query Match: 62.85% Indels: 8
DB: 8 Gaps: 2

US-10-051-909-36 (1-553) x AY093274 (1-1557)

QY 52 GluGluSerGlySerAspHisAspGlyValLeuArgArgProLeuLeuAsnThrGlySer 71
Db 25 GAAGAGGGAGAAATGAT-----CTTCGGCGACCGTCTCTACACACAGGAAGT 72
QY 72 TrpTyrArgMetSerSerArgGlnSerSerPheAlaProGlyThrSerSerMetAlaVal 91
Db 73 TGGTATCGGATGGGTTCGAGACAACTCTAGTATG-----TTGGAATCGTCTCAAGTT 123
QY 92 LeuArgGluSerHisValSerAlaPheLeuCysThrLeuLeValAlaLeuGlyProIle 111
Db 124 ATTCGAGACAGCTCAATTTCTGTCTTAGCTTGTGTTTGAATGTTGCTCTTGGTCTATT 183
QY 112 GlnPheGlyPheThrSerGlyPheSerSerProThrGlnAspAlaMetValArgAspLeu 131
Db 184 CAATTCGGATTCACTTGTGTTATTCTCTCACCACAACTCAAGTGCCATTACTAAGGATCTT 243
QY 132 AsnLeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyGlyMetVal 151
Db 244 GGTTTAACTGTATCAGAGTACTCTGTGTTTGGGTCTCTATCCCAATGTGGGTGCTATGTT 303
QY 152 GlyAlaIleAlaSerGlyGlnMetAlaGluTyrIleGlyArgLysGlySerLeuMetIle 171
Db 304 GGAGCAATTCCGAGTGGTTCAGATTGCTGAATACGTTGGGTCCTATCCCAATGTGGGTGCTATGTT 363
QY 172 AlaAlaIleProAsnIleIleGlyTrpLeuAlaIleSerPheAlaLysAspAlaSerPhe 191
Db 364 GCTGCAATTCCCAATATTATTGGATGGCTTTCGATATCATCTTGCAAAAGATACTCTTTT 423
QY 192 LeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTyrThrValPro 211
Db 424 CTTTACATGGGAAGATTGTTAGAGGCTTCGGAGTTGGGATTATCTCTATACGGTTCCC 483
QY 212 ValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGln 231
Db 484 GTATATATAGCTGAGATCGCTCCACAGACCATGAGAGAGCCCTTAGGTTTCAGTTAACCCAG 543
QY 232 LeuSerValThrPheGlyIlePheLeuAlaTyrLeuLeuGlyMetPheIleProTrpArg 251

Db 544 CTTTCTGTAAACAATTGGGATAATGCTGGCGTATTATTACTCGGTCTCTTTGTTCCATGGAGA 603
QY 252 LeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPhePheIle 271
Db 604 ATTCTTCAGATTTTGGGAGTATTGCCATGTACATTGTTGATACCGGGTCTTTTTCATT 663
QY 272 ProGluSerProArgTrpLeuAlaLysMetAsnLeuThrGluAspCysGluThrSerLeu 291
Db 664 CCTGAATCTCTCGTGGTGGTTGGCAAAGATGGGTTTGACAGATGATTTTCGAAACTTCATTG 723
QY 292 GlnValLeuArgGlyPheGluThrAspIleThrThrGluValAsnAspIleLysArgAla 311
Db 724 CAAGTTCTTCGTGGATTGAGACTGATATATCCGTTGAGGTTAATGAAATCAAGAGATCT 783
QY 312 ValAlaSerSerLysArgThrThrIleSerPheGlnGluLeuAsnGlnLysLysTyr 331
Db 784 GTGGCATCATCTAGCAAACGTTCTGCAGTTCGGTTTGTAGACCTCAAGCGCAGGAGATAC 843
QY 332 ArgThrProLeuLeuGlyIleGlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsn 351
Db 844 TATTTCCCACTGATGGTTGGTATAGGGCTGCTTGCACCTTCAACAACCTTGGAGGAATCAAT 903
QY 352 GlyValLeuPheTyrAlaSerSerIlePheLysAlaAlaGlyValThrAsnSerAspLeu 371
Db 904 GGTGTCTTGTCTATTTCGAGTACAATTTTGAATCTGCAGGGGTGACATCAAGTAATGTG 963
QY 372 AlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThrGlyValThrThrTrpLeuLeu 391
Db 964 GCGACATTTGGAGTTGGCGTTGTTCCAGGTAGTGGCGACTGGGATAGCGACATGGTTGGTG 1023
QY 392 AspArgAlaGlyArgArgIleLeuLeuIleIleSerThrSerGlyMetThrLeuCysLeu 411
Db 1024 GATAAAGCAGGTCGTTCGGCTTCTACTCATGATCTCTCTATCGGAATGACGATAGGCTC 1083
QY 412 LeuAlaValSerValPhePheLeuLysAspAsnIleSerGlnAspSerAsnSerTyr 431
Db 1084 GTGATTGTAGCAGTTGCATTTTACCTTAAGGAATTTGTATCACCTGATTCCACATGTAC 1143
QY 432 TyrIleLeuThrMetIleSerLeuValGlyIleValSerPheValIleThrPheSerPhe 451
Db 1144 AACATTCTAAGCATGGTTTCCGTAGTTGGAGTTGTGGCTATGGTTATTCTTGTCTCTCTA 1203
QY 452 GlyMetGlyAlaIleProTrpLeuMetMetSerGluIleLeuProValSerIleLysSer 471
Db 1204 GGAATGGGACCAATTCCATGGCTGATTATGTCTGAGATTCTGCCAGTGAATATAAAGGT 1263
QY 472 LeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeuThrSerPheAlaIleThrMetThr 491
Db 1264 TTAGCCGGAAGTATAGCAACTTTGTTAAACTGGTTCCGTGTCATGGTTAGTACAAATGACT 1323
QY 492 ThrAsnLeuMetLeuThrTrpSerValGlyGlyThrPheLeuSerTyrMetValValSer 511
Db 1324 GCAAATATGCTGTTAGCATGGAGCAGTGGGGAACTTTCACTCTCTACGCTTTGGTTGT 1383
QY 512 AlaPheThrIleValPheValValLeuTrpValProGluThrLysGly**AsnSerArg 531
Db 1384 GGATTCACCTGTGGTTTTTGTGAGTCTTTTGGGTTCTCTGAGACTAAAGGAAA-AACGCTTGA 1442
QY 532 GlyAspThrIlePheValSerLeu 539
Db 1443 AGAGATCCAAGCTTTGTTTCAGATG 1466

RESULT 6

AY059848 1866 bp mRNA linear PLN 04-NOV-2001
LOCUS Arabidopsis thaliana similar to integral membrane protein
DEFINITION (At1g19450; F18014.22) mRNA, complete cds.
ACCESSION AY059848
VERSION AY059848.1 GI:16648956
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis. 1 (bases 1 to 1866)

REFERENCE
AUTHORS
Nguyen,M., Karlin-Neumann,G., Southwick,A., Lam,B., Miranda,M., Palm,C.J., Bowser,L., Jones,T., Banh,J., Carninci,P., Chen,H., Cheuk,R., Chung,M.K., Hayashizaki,Y., Ishida,J., Kamiya,A., Kawai,J., Kim,C., Lin,J., Liu,S.X., Narusaka,M., Pham,P.K., Sakano,H., Sakurai,T., Satou,M., Seki,M., Shinn,P., Yamada,K., Shinozaki,K., Ecker,J., Theologis,A. and Davis,R.W.

TITLE
JOURNAL
Direct Submission
Submitted (22-OCT-2001) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA

COMMENT
e-mail for correspondence: arab@sequence.stanford.edu

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-Length cDNA'): Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Nguyen,M., Southwick,A., Karlin-Neumann,G., Lam,B., Miranda,M., Palm,C.J., Bowser,L., Jones,T., Banh,J., Chen,H., Cheuk,R., Chung,M.K., Kim,C., Lin,J., Liu,S.X., Pham,P.K., Sakano,H., Shinn,P., Yamada,K., Ecker,J., Theologis,A. and Davis,R.W.

Nguyen,M. (SSP/Stanford) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Davis,R.W. (SSP/Stanford) contributed equally to this work as PIs.

FEATURES
source

Location/Qualifiers
1..1866
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="I"
/clone="RAFL07-13-K02"
/note="This clone is in pBluescript"
1..1866
/gene="Atlg19450; F18O14.22"
249..1715
/gene="Atlg19450; F18O14.22"
/codon_start=1
/product="similar to integral membrane protein"
/protein_id="AAL24330.1"
/db_xref="GI:16648957"

gene
CDS

/translation="MSFRDDNTEGRNDRPFLHTGSMYRMGSRQSSMLSSQVIRD
SSISVLACVLIVALGPIQFGTSGYSSPTQAAITKDLGTVSEYSVFGSLSNVGMVG
AIASQIAEYVCRKGSMLIAAIPNIIGWLSISFAKDTSLYGRLLLEGFVGIIISYTV
PVYAEIAPOTMRGALGSVNOLSVTIGIMLAYLLGLFVPRILAVLGLPCTLLIPGL
FPFIPSPRLAKMGLTDDFETSLQVLRGFETDITVEVNEIKRSVASSSKRSARFVDL
KRRRYVFPIMVGIGLLAQOLGQINGVLFSYTSIFESAGVTSNNVATFGVGVQVVAT
GIATWLVDKAGRRLLLMISSIGMTISLVIVAVAFYLKEFVSPDSNMNINILSMVSVGV
VAMVISCSLGMGPPIPWLMSEILPVNIKLAGSIATLLNWFVSWLVTMTANMLLAWSS
GGTFTLYALVCGFTVVFVSLWVPETKKGKTLLEIQALFR"

ORIGIN

Alignment Scores:			
Pred. No.:	2.55e-121	Length:	1866
Score:	1746.50	Matches:	338
Percent Similarity:	83.40%	Conservative:	69
Best Local Similarity:	69.26%	Mismatches:	74
Query Match:	62.85%	Indels:	8
DB:	8	Gaps:	2

US-10-051-909-36 (1-553) x AY059848 (1-1866)

QY 52 GluGluserGlySerAspHisAspGlyValLeuArgArgProLeuLeuAsnThrGlySer 71
|||||
:::|
Db 273 GAAGAGGGAAGAAATGAT-----CTTCGGCAGCCGTTCTTACACACAGGAAGT 320
|||||
QY 72 TrpTyrArgMetSerSerArgGlnSerSerPheAlaProGlyThrSerSerMetAlaVal 91

Db 321 TGGTATCGGATGGGTTTCGAGACAATCTAGTATG-----TTGGAATCGTCTCAAGTT 371
|||||
QY 92 LeuArgGluSerHisValSerAlaPheLeuCysThrLeuIleValAlaLeuGlyProIle 111
:::|
Db 372 ATTCGAGACAGCTCAATTTCTGTCTTAGCTTGTGTTTGGTATTGCTCTGGTCTCTATT 431
|||||
QY 112 GlnPheGlyPheThrSerGlyPheSerSerProThrGlnAspAlaMetValArgAspLeu 131
|||||
Db 432 CAATTCGGATTCACTTGTGGTTATTCTCTCACCAACTCAAGCTGCCATTACTAAGGATCTT 491
|||||
QY 132 AsnLeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyGlyMetVal 151
|||||
Db 492 GGTTTAACTGTATCAGAGTACTCTGTGTTTGGGTCTCTATCCAATGTGGTGTATGGTT 551
|||||
QY 152 GlyAlaIleAlaSerGlyGlnMetAlaGluTyrIleGlyArgLysGlySerLeuMetIle 171
|||||
Db 552 GGAGCAATTGCGAGTGGTCAGATTGCTGAATACGTGGACGGAAAGGTCCTCTGATGATT 611
|||||
QY 172 AlaAlaIleProAsnIleIleGlyTrpLeuAlaIleSerPheAlaLysAspAlaSerPhe 191
|||||
Db 612 GCTGCAATTCCTCAATATATTGGATGGCTTTTCGATATCATTTGCAAAAGATACTTCTTT 671
|||||
QY 192 LeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTyrThrValPro 211
|||||
Db 672 CTTTACATGGGAAGATTGTTAGAGGCTTCGGAGTTGGGATTATCTCCTATACGGTTCCC 731
|||||
QY 212 ValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGln 231
|||||
Db 732 GTATATATAGCTGAGATCGCTCCACAGACCATTGAGAGGAGCCTTAGGTTTCAGTTAACCCAG 791
|||||
QY 232 LeuSerValThrPheGlyIlePheLeuAlaTyrLeuLeuGlyMetPheIleProTyrArg 251
|||||
Db 792 CTTTCTGTAAACAATTGGGATAAATGCTGGCGTATTACTCGGTCTCTTTTGTTCATGGAGA 851
|||||
QY 252 LeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPhePheIle 271
:::|
Db 852 ATTCTTGCAGTTTGGGAGTATTGCCATGTACATTGTTGATACCGGGTCTTTTTTTCATT 911
|||||
QY 272 ProGluSerProArgTrpLeuAlaLysMetAsnLeuThrGluAspCysGluThrSerLeu 291
|||||
Db 912 CCTGAATCTCCTCGTGGTGTGGCAAAGATGGGTTTGCACAGATGATTTTCGAAACTTCATTG 971
|||||
QY 292 GlnValLeuArgGlyPheGluThrAspIleThrThrGluValAsnAspIleLysArgAla 311
|||||
Db 972 CAGTTCTTCGTGGATTTCGAGACTGATATTACCGTTGAGGTTAATGAAATCAAGAGATCT 1031
|||||
QY 312 ValAlaSerSerSerLysArgThrThrIleSerPheGlnGluLeuAsnGlnLysLysTyr 331
|||||
Db 1032 GTGCGATCATCTAGCAAACGTTCTGCAGTTCGGTTTGTAGACCTCAAGCGCAGGAGATAC 1091
|||||
QY 332 ArgThrProLeuLeuGlyIleGlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsn 351
|||||
Db 1092 TATTTCCCACTGATGGTTGGTATAGGGCTGCTTGCACTTCAACAACCTTGGAGGAATCAAT 1151
|||||
QY 352 GlyValLeuPheTyrAlaSerSerIlePheLysAlaAlaGlyValThrAsnSerAspLeu 371
|||||
Db 1152 GGTGCTTGTCTATTTCGAGTACAATTTTGAATCTGCAGGGGTGACATCAAGTAATGTG 1211
|||||
QY 372 AlaThrCysSerLeuGlyValAlaIleGlnValLeuAlaThrGlyValThrThrTrpLeuLeu 391
|||||
Db 1212 GCGACATTTGGAGTTGGCGTGTTCAGGTAGTGGCGACTGGGATAGCGACATGGTTGGTG 1271
|||||
QY 392 AspArgAlaGlyArgArgIleLeuLeuIleIleSerThrSerGlyMetThrLeuCysLeu 411
|||||
Db 1272 GATAAGCAGGTCGTTCGGCTTCTACTCATGATCTCTTCTATCGGAATGACGATAAGCCTC 1331
|||||
QY 412 LeuAlaValSerValValPhePheLeuLysAspAsnIleSerGlnAspSerAsnSerTyr 431
:::|
Db 1332 GTGATTGTAGCAGTTGCATTTTACCTTAAGGAATTTGTATCACCTGATTCCCAACATGTAC 1391
|||||
QY 432 TyrIleLeuThrMetIleSerLeuValGlyIleValSerPheValIleThrPheSerPhe 451
|||||
:::|

Db 1392 AACATTCTAAGCATGGTTTCCGCTAGTTGGAGTTGTGGCTATGCTATTCTTCTCTCTA 1451
QY 452 GlyMetGlyAlaIleProTirPLeuMetMetSerGluIleLeuProValSerIleLysSer 471
Db 1452 GGAATGGGACCAATTCCATGGCTGATATGCTGAGATTCTGCAGTGAATATAAGGGT 1511
QY 472 LeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeuThrSerPheAlaIleThrMetThr 491
Db 1512 TTAGCCGGAAGTATAGCAACTTTGTAAACTGGTTCTGTCATGTTAGTGACATGACT 1571
QY 492 ThrAsnLeuMetLeuThrTrpSerValGlyGlyThrPheLeuSerTyrMetValValSer 511
Db 1572 GCAAATATGCTGTAGCATGGAGCAGTGGGGAACTTTCACTCTCTACGCTTTGGTTGT 1631
QY 512 AlaPheThrIleValPheValValLeuTrpValProGluThrLysGly**AsnSerArg 531
Db 1632 GGATTCACCTGTGGTTTGTGAGCTTTGGGTTCTCTGAGACTAAAGGAAA-AACGCTGA 1690
QY 532 GlyAspThrIlePheValSerLeu 539
Db 1691 AGAGATCCAGCTTTGTTTCAGATG 1714

RESULT 7
AF412060 1695 bp mRNA linear PLN 24-SEP-2001
LOCUS Arabidopsis thaliana At1g75220/F22H5_6 mRNA, complete cds.
DEFINITION AF412060
ACCESSION AF412060
VERSION AF412060.1 GI:15724239
KEYWORDS FLI-CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 1695)
Cheuk,R., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Banh,J.,
Bowser,L., Carninci,P., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y.,
Ishida,J., Jiang,P.X., Jones,T., Kamiya,A., Karlin-Neumann,G.,
Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M.,
Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K.,
Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C.,
Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K.,
Davis,R.W., Theologis,A. and Ecker,J.R.
Arabidopsis cdna clones
Unpublished

2 (bases 1 to 1695)
Cheuk,R., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Banh,J.,
Bowser,L., Carninci,P., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y.,
Ishida,J., Jiang,P.X., Jones,T., Kamiya,A., Karlin-Neumann,G.,
Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M.,
Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K.,
Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C.,
Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K.,
Davis,R.W., Theologis,A. and Ecker,J.R.
Direct Submission
Submitted (21-AUG-2001) Salk Institute Genomic Analysis Laboratory
(SIGNAL), Plant Biology Laboratory, The Salk Institute for
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
USA
RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
Arabidopsis Full-Length cDNA') : Seki,M., Narusaka,M., Ishida,J.,
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGECC (SSP) Consortium members carried out the
sequencing and annotation of the RAFL cDNAs: Cheuk,R., Chen,H.,
Kim,C.J., Koesema,E., Meyers,M.C., Shinn,P., Banh,J., Bowser,L.,
Dale,J.M., Goldsmith,A.D., Jiang,P.X., Jones,T., Karlin-Neumann,G.,
Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M., Nguyen,M.,
Onodera,C.S., Palm,C.J., Pham,P.K., Quach,H.L., Southwick,A.,
Tang,C.C., Toriumi,M., Yamada,K., Yamamura, Y., Yu,G., Yu,S.,
Davis,R.W., Theologis,A., and Ecker,J.R.

Cheuk,R. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to
this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk)
contributed equally to this work as PIs.

FEATURES
source

1. .1695
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="1"
/clone="RAFL06-75-L22(R13497)"
/ecotype="Columbia"
1. .11
12. .1475
/note="putative integral membrane protein"
/codon_start=1
/product="At1g75220/F22H5_6"
/protein_id="AAL06513.1"
/db_xref="GI:15724240"
/translation="MSFRDDNEEARNDLRRPFHTGWSYRMGSRQSSMMGSSQVIRDS
SISVLACVLI VALGPIQFGFTCGYSSPTQAAITKDLTVSEYSVFCSLSNVGAMVGA
IASGQIAEYIGRKGLMAAIPNIIGWLCISFAKDTSLYMGRLLEGFGVGIISYTVP
VYIAIAPQNMRRGGLSVNQLSVTIGIMLAYLLGLFVPRILAVLGLPCTLLIPGLF
PIESPRLAKMGMTDEFETSLQVRGFETITVEWEIKRSVASSTKNTVRFVDLK
RRRYFPLMVGIGLLVQLGGINGLVFYSSTIFESAGVTSSNAATFGVCAIQVVATA
ISTWLVDKAGRRLLTISVGMTISLVI VAAAFYLKEFVSPDSMDYSLSLSVVGWV
AMVVFSLGMGPIPWIMSEILPVNIKGLAGSIATLANWFFSWLITMTANLLAWSSG
GTFTLYGLVCAFTVVFVTLWVPETKGTLEELQSLFR"
1476. .1695

5'UTR
CDS

3'UTR
ORIGIN

Alignment Scores:
Pred. No.: 8.87e-121 Length: 1695
Score: 1738.50 Matches: 343
Percent Similarity: 82.47% Conservative: 71
Best Local Similarity: 68.33% Mismatches: 81
Query Match: 62.56% Indels: 8
DB: 2
US-10-051-909-36 (1-553) x AF412060 (1-1695)
QY 52 GluGluSerGlySerAspHisAspGlyValLeuArgArgProLeuLeuAsnThrGlySer 71
Db 33 GAGGAGGCTAGGAATGAT-----TTGGCGTCGGCGGTTTATTACACACCGGAAGC 80
QY 72 TrpTyrArgMetSerSerArgGlnSerSerPheAlaProGlyThrSerSerMetAlaVal 91
Db 81 TGGTACCGGATGGGTTTCGAGACAATCGAGTATG-----ATGGGTTCTCTCAAGTT 131
QY 92 LeuArgGluSerHisValSerAlaPheLeuCysThrLeuIleValAlaLeuGlyProIle 111
Db 132 ATTCGAGACAGCTCTATCTCTGTTCTTTCATGATGTTGATCGTTGCTTGGCCCTATT 191
QY 112 GlnPheGlyPheThrSerGlyPheSerSerProThrGlnAspAlaMetValArgAspLeu 131
Db 192 CAATTCCGATTCACCTTGTGTTTATTCATCTCCAACTCAAGCTGCAATCACCAGGATCTC 251
QY 132 AsnLeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyGlyMetVal 151
Db 252 GGTTTAAACAGTATCTGAGTACTCTGTGTTTGGTTCTCTTCCAATGTGGTGCTATGTT 311
QY 152 GlyAlaIleAlaSerGlyGlnMetAlaGluTyrIleGlyArgLysGlySerLeuMetIle 171
Db 312 CGTGCAATTGCCAGTGGTCAGATTGCAGATACATTGGACGAAAGGGTCTCTTATGATT 371
QY 172 AlaAlaIleProAsnIleIleGlyTrpLeuAlaIleSerPheAlaLysAspAlaSerPhe 191
Db 372 GCTGCAATTCCCAATATCATCTGATGGCTTTGCAATATCATTTGCAAAAGATACTTCTTT 431
QY 192 LeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTyrThrValPro 211
Db 432 CTTTACATGGGAAGACTTTTAGAAGGCTTTGGCGTTGGGATCATCTCTTACACTGTACCT 491

TITLE	Arabidopsis cDNA clones		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 1914)		
AUTHORS	Cheuk,R., Chen,H., Kim,C.J., Meyers,M.C., Banh,J., Bowser,L., Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K., Yamamura,Y., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.		
TITLE	Direct Submission		
JOURNAL	Submitted (27-DEC-2001) Salk Institute Genomic Analysis Laboratory (SIGNAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA		
COMMENT	RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-length cDNA') : Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.		
FEATURES	The Salk, Stanford, PGEN (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Cheuk,R., Chen,H., Kim,C.J., Meyers,M.C., Shinn,P., Banh,J., Bowser,L., Dale,J.M., Goldsmith,A.D., Jiang,P.X., Jones,T., Karlin-Neumann,G., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K., Quach,H.L., Southwick,A., Tang,C.C., Toriumi,M., Yamada,K., Yamamura, Y., Yu,G., Yu,S., Davis,R.W., Theologis,A., and Ecker,J.R.		
source	Cheuk,R. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk) contributed equally to this work as PIs.		
Location/Qualifiers	1. .1914		
source	/organism="Arabidopsis thaliana"		
	/mol_type="mRNA"		
	/db_xref="taxon:3702"		
	/chromosome="1"		
	/clone="RAFL09-48-G03 (R21488)"		
	/ecotype="Columbia"		
ORIGIN			
Alignment Scores:			
Pred. No.:	8.32e-120	Length:	1914
Score:	1726.50	Matches:	337
Percent Similarity:	83.03%	Conservative:	69
Best Local Similarity:	68.92%	Mismatches:	75
Query Match:	62.13%	Indels:	9
DB:	8	Gaps:	2
US-10-051-909-36 (1-553) x AF462803 (1-1914)			
QY	52	GluGluSerGlySerAspHisAspGlyValLeuArgArgProLeuLeuAsnThrGlySer	71
Db	274	GAAGAGGGAAGAAATGAT-----CTTCGGCGACCGTTCTTACACACAGGAAGT	321
QY	72	TrpTyrArgMetSerSerArgGlnSerSerPheAlaProGlyThrSerSerMetAlaVal	91
Db	322	TGGTATCGGATGGGTTTCGAGACAAATCTAGTAIG-----TTGGAATCGTCTCAAGTT	372
QY	92	LeuArgGluSerHisValSerAlaPheLeuCysThrLeuIleValAlaLeuGlyProIle	111
Db	373	ATTCGAGACAGCTCAATTTCTGCTCTTAGCTTGTGTTTGTGATTGTTGCTCTTGGTCCTATT	432
QY	112	GlnPheGlyPheThrSerGlyPheSerSerProThrGlnAspAlaMetValArgAspLeu	131
Db	433	CAATTGGATTCACTTGTGGTTATTCTTCACCAACTCAAGCTGCCATTACTAAGGATCTT	492
QY	132	AsnLeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyGlyMetVal	151
Db	493	GGTTTAACGTATCAGAGTACTCTGTGTTGGTCTCTATCCAATGTGGGTGTATGGTT	552
QY	152	GlyAlaIleAlaSerGlyGlnMetAlaGluTyrIleGly-ArgLysGlySerLeuMetIl	171
Db	553	GGAGCAATTGGAGTGGTCAGATTGCTGAATACGTTGGACTTGAAAGGGTCTCTGATGAT	612
QY	171	eAlaAlaIleProAsnIleIleGlyTrpLeuAlaIleSerPheAlaLysAspAlaSerPh	191
Db	613	TGCTGCAATTCCCAATATTATTGGATGGCTTTCGATATCATTTGCAAAAGATACTTCTTT	672
QY	191	eLeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTyrThrValPr	211
Db	673	TCTTTACATGGGAAGATTGTTAGAAGGCTTCGGAGTTGGGATTATCTCCTATACGGTTCC	732
QY	211	oValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGl	231
Db	733	CGTATATATAGCTGAGATCGCTCCACAGACCATTGAGAGGACCTTAGGTTCAAGTTAACCA	792
QY	231	nLeuSerValThrPheGlyIlePheLeuAlaTyrLeuLeuGlyMetPheIleProTrpAr	251
Db	793	GCTTCTGTAAACAATTGGGATAAATGCTGGCGTAATTACTCGGTCTCTTTGTTCCATGGAG	852
QY	251	gLeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPhePheIl	271
Db	853	AATTCTTGCAATTTTGGAGTATTGCCATGTACATTTGATACCGGGTCTTTTTCAT	912
QY	271	eProGluSerProArgTrpLeuAlaLysMetAsnLeuThrGluAspCysGluThrSerLe	291
Db	913	TCCTGAATCTCCTCGTGGTGGCAAGATGGGTTTGACAGATGATTTCGAAACTTCATT	972
QY	291	uGlnValLeuArgGlyPheGluThrAspIleThrThrGluValAsnAspIleLysArgAl	311
Db	973	GCAAGTTCTTCGTGGATTGAGACTGATATTACCGTTGAGGTTAATGAAATCAAGAGATC	1032
QY	311	aValAlaSerSerSerLysArgThrThrIleSerPheGlnGluLeuAsnGlnLysLysTy	331
Db	1033	TGTGGCATCATCTAGCAACGTTCTGCAGTTCGGTTTGTAGACCTCAAGCGCAGGAGATA	1092
QY	331	rArgThrProLeuLeuGlyIleGlyLeuLeuValLeuGlnAsnLeuSerGlyIleAs	351
Db	1093	CTATTCCCACCTGATGGTGGTATAGGCTGCTTGCACCTTCAACAACCTTGGAGGAATCAA	1152
QY	351	nGlyValLeuPheTyrAlaSerSerIlePheLysAlaAlaGlyValThrAsnSerAspLe	371
Db	1153	TGGTGTCTTGTCTATTTCGAGTACAAATTTTGAATCTCGAGGGGTGACATCAAGTAATGT	1212
QY	371	uAlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThrGlyValThrThrTrpLeuLe	391
Db	1213	GGCGACATTTGGAGTTGGCGTTGTTTCAGGTAGTGGCGACCTGGGATAGCGACATGGTTGGT	1272
QY	391	uAspArgAlaGlyArgArgIleLeuLeuIleIleSerThrSerGlyMetThrLeuCysLe	411
Db	1273	GGATAAAGCAGGTCGTCGGCTTCTACTCATGATCTCTCTATCGGAATGACGATAAGCCT	1332
QY	411	uLeuAlaValSerValPhePheLeuLysAspAsnIleSerGlnAspSerAsnSerTy	431
Db	1333	CGTGATTGTAGCAGTTGCATTTTACCTTAAGGAATTTGTATCACCTGATTCACACATGTA	1392
QY	431	rTyrIleLeuThrMetIleSerLeuValGlyIleValSerPheValIleThrPheSerPh	451
Db	1393	CAACATTCTAAGCATGGTTTCGGTAGTTGGAGTTGTGGCTATGGTTATTCTTGCTCTCT	1452
QY	451	eGlyMetGlyAlaIleProTrpLeuMetMetSerGluIleLeuProValSerIleLysSe	471
Db	1453	AGGAATGGACCAATTCCATGGCTGATTATGTCTGAGATTCGCCAGTGAATATAAAGG	1512
QY	471	rLeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeuThrSerPheAlaIleThrMetTh	491
Db	1513	TTTAGCCGGAAGTATAGCAACTTTGTTAAACTGGTTCCGTGTCATGGTTAGTGACAATGAC	1572
QY	491	rThrAsnLeuMetLeuThrTrpSerValGlyGlyThrPheLeuSerTyrMetValValse	511
Db	1573	TGCAATATGCTGTTAGCATGGAGCAGTGGGGGAACCTTTCATCTCTACGCTTTGGTTG	1632


```

RESULT 11
BT014234
LOCUS
DEFINITION
ACCESSION
BT014234
VERSION
BT014234.1
KEYWORDS
FLI_CDNA
SOURCE
Lycopersicon esculentum (tomato)
ORGANISM
Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1
(bases 1 to 1825)
REFERENCE
Kirkness,E.F., Wang,W. and Vazeille,A.
Direct Submission
Submitted (11-MAY-2004) The Institute for Genomic Research, 9712
Medical Center Drive, Rockville, MD 20850, USA
FEATURES
Location/Qualifiers
1..1825
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/db_xref="taxon:4081"
/clone="133433R"
/tissue type="fruit"
/note="TZXHM68"
ORIGIN
Alignment Scores:
Pred. No.: 2,52e-117 Length: 1825
Score: 1693.00 Matches: 328
Percent Similarity: 82.85% Conservative: 73
Best Local Similarity: 67.77% Mismatches: 79
Query Match: 60.92% Indels: 5
DB: 8 Gaps: 2
US-10-051-909-36 (1-553) x BT014234 (1-1825)
QY 57 AspHisAspGlyValLeuArgArgProLeuLeuAsnThrGlySerTyrTyrArgMetSer 76
Db 122 AACAAATGAAGGGATCTGAAAAGCCATTTTACATACTGGGAGTTGGTATCGTATGGG 181
QY 77 SerArgGlnSerSerPheAlaProGlyThrSerSerMetAlaValLeuArgGluSerHis 96
Db 182 TCAAGACAGTCTAGTAGT-----ATGGGTTCTTCTCAAGCTATTCTGTGATGTTCT 232
QY 97 ValSerAlaPheLeuCysThrLeuIleValAlaLeuGlyProIleGlnPheGlyPheThr 116
Db 233 ATTCTGTCTTGTGTTGTCTCATAGTTGCTTTGGGCCCCAATCCAGTTTGGTTTCACT 292
QY 117 SerGlyPheSerSerProThrGlnAspAlaMetValArgAspLeuAsnLeuSerIleSer 136
Db 293 TGTGGCTACACTTCACCAACGCAAAATTCGAATTAACCAAGATCTGAAGCTCACTGTTTCT 352
QY 137 GluPheSerAlaPheGlySerLeuSerAsnValGlyGlyMetValGlyAlaIleAlaSer 156
Db 353 GAGTCTCTTGTTCGGTCTTATCCAAATGTGGGTGCTATGTTGGTGCAATATCTAGT 412
QY 157 GlyGlnMetAlaGluTyrIleGlyArgLysGlySerLeuMetIleAlaAlaIleProAsn 176
Db 413 GGTCAAATTTCTGAGTACATTGGCGGAAAAGGGTCTTTAATGATAGTGTGTCATACCAAA 472
QY 177 IleIleGlyTrpLeuAlaIleSerPheAlaLysAspAlaSerPheLeuTyrMetGlyArg 196
Db 473 ATCATTTGGTTGGCTTGCCATCTCATTTGCCAAAGATCCTTCTTTCTATACATGGGAAGA 532
QY 197 LeuLeuGluGlyPheGlyValGlyIleIleSerTyrThrValProValTyrIleAlaGlu 216
Db 533 TTGTTGGAGGGCTTTGGTGTGGGAATAATCTCTTATGTGGTTCCTGTATATATGCTGAG 592
QY 217 IleSerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPhe 236
Db 593 ATTGCACCTCAGAAATATGAGAGGGGGTCTAGGGTCTGTTTAACCAAGCTCTCTGTACAATT 652
```

```

QY 237 GlyIlePheLeuAlaTyrLeuLeuGlyMetPheIleProTyrPArgLeuLeuAlaValIle 256
Db 653 GGGATCATGTGTCTCTATTACTGGGACTTTTGTTAATTTGGAGGCTGCTGTGTTCTT 712
QY 257 GlyAlaLeuProCysThrMetLeuIleProGlyLeuPhePheIleProGluSerProArg 276
Db 713 GGAGCATTTACCTTGCCTAGCACGTGATACCTGGACTATTTTTCATCCCGAATCTCCTCG 772
QY 277 TrpLeuAlaLysMetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuArgGly 296
Db 773 TGGTTGGCCAAATGGGTTTAAACAGAAGATTTTGAACACTTCGTTGCAAGTCTTCGAGGC 832
QY 297 PheGluThrAspIleThrThrGluValAsnAspIleLysArgAlaValAlaSerSer--- 315
Db 833 TTTGAGGCTGATATTACCAATGAAGTAATGAATTAAGAGGTCTGTAGCATCATCAACA 892
QY 316 SerLysArgThrThrIleSerPheGlnGluLeuAsnGlnLysLysTyrArgThrProLeu 335
Db 893 AACCGAAGAACAGCAATGCGTTTTCGGGATCTCAACATAGAAAGATATTGGTTCTCTC 952
QY 336 LeuLeuGlyIleGlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPhe 355
Db 953 ATGGTAGGAATTGGACTACTTATCTCTGCAACAGTAGGAGGAATCAACGGTGTATATTC 1012
QY 356 TyrAlaSerSerIlePheLysAlaAlaGlyValThrAsnSerAspLeuAlaThrCysSer 375
Db 1013 TATTCAGCAACATTTTCTCTATCTGCGGTATTTCGTCAGTAATGCTGCAACTTTTGGT 1072
QY 376 LeuGlyAlaIleGlnValLeuAlaThrGlyValThrTrpLeuLeuAspArgAlaGly 395
Db 1073 GTCGGTGCTATACAGGTTGTGCAACTGCAGTTTCTACATGGTGGTGGACAGAACTGGA 1132
QY 396 ArgArgIleLeuLeuIleIleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSer 415
Db 1133 CGTAGAATTTTATTAATTGTTTCTCGGCTGGAATGGCTGTAGTCTTCTTATTGTCTCC 1192
QY 416 ValValPhePheLeuLysAspAsnIleSerGlnAspSerAsnSerTyrTyrIleLeuThr 435
Db 1193 GTTGCTTCTATCTAAAGGTTATGTATCTCTGAGGATTTCTACCTCTATGGTGTCTTGGGA 1252
QY 436 MetIleSerLeuValGlyIleValSerPheValIleThrPheSerPheGlyMetGlyAla 455
Db 1253 ATAGTGTCCGTAGTTGGGGTGTGTTTCATGATTATTTCATTCTCGCTCGGAATGGGACCA 1312
QY 456 IleProTrpLeuMetMetSerGluIleLeuProValSerIleLysSerLeuGlyGlySer 475
Db 1313 ATTCCATGGCTTATAATGTCTGAGATTCTTCCAATCAAGATCAAAAGGTCTCGCTGGAAGT 1372
QY 476 IleAlaThrLeuAlaAsnTrpLeuThrSerPheAlaIleThrMetThrThrAsnLeuMet 495
Db 1373 ACGGCAACATTAGCCAACTGGTTATTCTCGTGGGTAATAACAGTGACTGCACCGTTGCTT 1432
QY 496 LeuThrTrpSerValGlyGlyThrPheLeuSerTyrMetValSerAlaPheThrIle 515
Db 1433 ATGGCTTGAGCAGTGGAGGAACCTTCACTATTATCATGTTATGTGTGTTTCACCGTG 1492
QY 516 ValPheValValLeuTrpValProGluThrLysGly**AsnSerArgGlyAspThrIle 535
Db 1493 GCGTTTCGTGCAATTTGGTACCAGAGACGAAAGGGAG-AACTCTAGAGGAAATTCAGTT 1551
QY 536 PheValSerLeu 539
Db 1552 TTCTTTTCAGATG 1563
RESULT 12
AX653120
LOCUS
DEFINITION
Sequence 2990 from Patent WO03000898.
ACCESSION
AX653120
VERSION
AX653120.1
KEYWORDS
GI:29155934
Oryza sativa
```

AX653120 1428 bp DNA linear PAT 22-MAR-2003

ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.

REFERENCE 1
AUTHORS Chang, H.S., Chen, W., Cooper, B., Glazebrook, J., Goff, S.A., Hou, Y.M.,
Katagiri, F., Quan, S., Tao, Y., Whitham, S., Xie, Z., Zhu, T. and Zou, G.
TITLE Plant genes involved in defense against pathogens
JOURNAL Patent: WO 0300898-A 2990 03-JAN-2003;
SYNGENTA PARTICIPATIONS AG (CH)

FEATURES
source Location/Qualifiers
1. .1428
/organism="Oryza sativa"
/mol_type="unassigned DNA"
/db_xref="taxon:4530"

ORIGIN
Alignment Scores:
Pred. No.: 8.83e-109 Length: 1428
Score: 1577.00 Matches: 318
Percent Similarity: 85.85% Conservative: 34
Best Local Similarity: 77.56% Mismatches: 32
Query Match: 56.75% Indels: 27
DB: 6 Gaps: 2

US-10-051-909-36 (1-553) x AX653120 (1-1428)

QY 130 AspLeuAsnLeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyGly 149
DB 277 GATCTGTACITACAAACAGCATATGTTCTCGGCTTTCGGTTCCGTGTCACACGTCGGCC 336

QY 150 MetValGlyAlaIleAlaSerGlyGlnMetAlaGluTyrIleGlyArgLysGlySerLeu 169
DB 337 ATGGTCGGAGCGATCGCCAGTGGCGAGATGGCGGAGTACATGGCCGGAAGGGTCTTG 396

QY 170 MetIleAlaAlaIleProAsnIleIleGlyTyrLeuAlaIleSerPheAlaLysAspAla 189
DB 397 ATAATTGCGCGGTTCCTAACATCATGTTGGTTGGCTTGCCTTCCTTTGCAAAAGACGCG 456

QY 190 SerPheLeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTyrThr 209
DB 457 TCATTTTATACATGGGACGCTTGTGTAAGGTTTGTGGTGTGTTGTTATATATATACG 516

QY 210 ValProValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerVal 229
DB 517 GTGCCAGTATACATAGCAGAGATATCTCATCAGAACACAGAGGAGCACATTTGGCTCCGTG 576

QY 230 AsnGlnLeuSerValThrPheGlyIlePheLeuAlaTyrLeuLeuGlyMetPheIlePro 249
DB 577 AACCAAGTGTCCGTTACCATTTGCTATCTTGTGGCCTATTTGCTAGGCATGTTTGTCTCT 636

QY 250 TrpArgLeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPhe 269
DB 637 TGGAGGCTGCTTGCAGTATAGGAAGCATCCCATGTATCATTTGTTAATACCTGCTCTATTC 696

QY 270 PheIleProGluSerProArgTyrLeuAlaLysMetAsnLeuThrGluAspCysGluThr 289
DB 697 TTCAATCCGGAATCCCAAGATGGCTGGCAAAATGAAATGATGGATGATTTGAGGCT 756

QY 290 SerLeuGlnValLeuArgGlyPheGluThrAspIleThrThrGluValAsnAspIleLys 309
DB 757 TCTCTACAAAGTTTGGGGGTTTGAGACTGATATTAATGTCAGAA----- 801

QY 310 ArgAlaValAlaSerSerSerLysArgThrThrIleSerPheGlnGluLeuAsnGlnLys 329
DB 802 AGACAGTAGCATCGGCAACAAAGAACACACAGTCCGCTTTTAAAGAGTTGAACCAAAAG 861

QY 330 LysTyrArgThrProLeuLeuLeuGlyIleGlyLeuLeuValLeuGlnAsnLeuSerGly 349
DB 862 AAATACCGCACTCCCTTACTGATAGGAAGTGGCTTCTTGTACTTTCAGAAATCTAAGTGA 921

QY 350 IleAsnGlyValLeuPheTyrAlaSerSerIlePheLysAlaAlaGlyValThrAsnSer 369

DB 922 ATAAATGGTATTCTGTTTATGCAAGTAGAATCTTTCAGAGATGCAGGGTTTACAAACAGT 981

QY 370 AspLeuAlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThrGlyValThrTyr 389
DB 982 GACTTGGCCACATGTCACACTAGGAGCAATTCAG----- 1014

QY 390 LeuLeuAspArgAlaGlyArgArgIleLeuLeuLeuIleSerThrSerGlyMetThrLeu 409
DB 1015 -----ATCTCTACTGCTGGGATGACTCTA 1038

QY 410 CysLeuLeuAlaValSerValPhePheLeuLysAspAsnIleSerGlnAspSerAsn 429
DB 1039 AGCCTTCTTGCAGTATCTGTTGTTATTTTCTGGAGGGTAATATTTTTCATGATTCTCAT 1098

QY 430 SerTyrTyrIleLeuThrMetIleSerLeuValGlyIleValSerPheValIleThrPhe 449
DB 1099 TCGTTCTACATCTTAAGTATGATCTCCTTGGTGTCTTGTGGCTTATATCATCACCTTT 1158

QY 450 SerPheGlyMetGlyAlaIleProTyrLeuMetMetSerGluIleLeuProValSerIle 469
DB 1159 TCCTTCGGCATGGGTGCCATTCATGGGTGATGATGCTGAGATCTCCAGTTAGCATC 1218

QY 470 LysSerLeuGlyGlySerIleAlaThrLeuAlaAsnTyrLeuThrSerPheAlaIleThr 489
DB 1219 AAGAGTCTTGGGGGAAGCTTTGCGACACTTGCACATGCTTACATCTCTGGCAATAACA 1278

QY 490 MetThrThrAsnLeuMetLeuThrTyrSerValGlyGlyThrPheLeuSerTyrMetVal 509
DB 1279 ATGACAGCAATTTGTTACTCAGCTGGAGTGTCTGGAGGAACATTTTGTCTACATGATT 1338

QY 510 ValSerAlaPheThrIleValPheValValLeuTyrValProGluThrLysGly***Asn 529
DB 1339 GTGAGTGTCTTCACTCTCGTGTTCGTATATTTTGGTGGCAGAGACAAAGGAAG-AAC 1397

QY 530 SerArgGlyAspThrIlePheValSerLeu 539
DB 1398 TCTGGAGGAGATACAGTTTTCGTTCTGCTA 1427

RESULT 13
AC120986 204649 bp DNA linear PLN 23-APR-2004
LOCUS Oryza sativa (japonica cultivar-group) chromosome 5 clone
DEFINITION OJ1781_H11, complete sequence.
AC120986
AC120986.2 GI:46518554
KEYWORDS HTG.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (bases 1 to 204649)
AUTHORS Chow, T.-Y., Hsing, Y.-I.C., Chen, C.-S., Chen, H.-H., Liu, S.-M.,
Chao, Y.-T., Chang, S.-J., Chen, H.-C., Chen, S.-K., Chen, T.-R.,
Chen, Y.-L., Cheng, C.-H., Chung, C.-I., Han, S.-Y., Hsiao, S.-H.,
Hsiung, J.-N., Hsu, C.-H., Huang, J.-J., Kau, P.-I., Lee, M.-C.,
Leu, H.-L., Li, Y.-F., Lin, S.-J., Lin, Y.-C., Wu, S.-W., Yu, C.-Y.,
Yu, S.-W., Wu, H.-P. and Shaw, J.-F.
TITLE Oryza sativa BAC OJ1781_H11 genomic sequence
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 204649)
AUTHORS Chow, T.-Y. and Hsing, Y.-I.C.
TITLE Direct Submission
JOURNAL Submitted (16-MAY-2002) Institute of Botany, Academia Sinica, 128,
Section 2, Academia Road, Nankang, Taipei 11529, Taiwan
REFERENCE 3 (bases 1 to 204649)
AUTHORS Hsing, Y.-I.C. and Chow, T.-Y.
TITLE Direct Submission
JOURNAL Submitted (29-MAR-2003) Institute of Botany, Academia Sinica, 128,
Section 2, Yien-chu-yuan Road, Nankang, Taipei 11529, Taiwan
REFERENCE 4 (bases 1 to 204649)
AUTHORS Chow, T.-Y.
TITLE Direct Submission

JOURNAL Submitted (23-APR-2004) Institute of Botany, Academia Sinica, 128, Section 2, Academia Road, Nankang, Taipei 11529, Taiwan
COMMENT On Apr 23, 2004 this sequence version replaced gi:20806177.
The nucleotide sequence of this BAC clone was generated by combining Monsanto and ASPGC-Taiwan sequencing data.

FEATURES
source
1. .204649
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="5"
/clone="OJ1781_H11"

ORIGIN
Alignment Scores:
Pred. No.: 1.58e-96 Length: 204649
Score: 1453.50 Matches: 426
Percent Similarity: 37.12% Conservative: 45
Best Local Similarity: 33.57% Mismatches: 54
Query Match: 52.30% Indels: 747
DB: 8 Gaps: 19

US-10-051-909-36 (1-553) x AC120986 (1-204649)

QY 3 SerSerSerSerPheArgProAlaGlyLysLysLysLysLysAsnGlnGlyLeuArg 22
Db 118274 TCTCTCTCATCCATTCTCCGGTTGAGCGGCGCGC----- 118309

QY 23 ArgGluAlaValProGlyArgProAlaSerGluLeuArgThrArgValMetGlyGly 42
Db 118310 -----GGCGGAGCGGAGCGAGCGGCGGCGGATGAACGGCGCGGG 118357

QY 43 SerAsnArgGlyGlyAlaGlyAlaGlyGluGluSerGlySerAspHisAsp----- 59
Db 118358 AGCAGGCGGGAGGTGCGTCGTGCGCGCAGAGCGGCGGAGCGGCGGCGGCGCAGC 118417

QY 60 ---GlyValLeuArgArgProLeuLeuAsnThrGlySerTyrArgMetSerSerArg 78
Db 118418 GCGGCGGCGTTCGGGAAGCCGCTGCTGAACACGGGAGCTGGTACAGATGGATCGCGG 118477

QY 79 GlnSerSerPheAlaProGlyThrSerSerMetAlaValLeuArgGluSerHisValSer 98
Db 118478 TCCAGCCTCGCGCC-----TCCTCCATGGCGGCCCATCCGGGAGTCCCACGTCTCC 118528

QY 99 AlaPheLeuCysThrLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGly 118
Db 118529 GCCTTCTCTGCACGCTCATCGTCGGCTCGGCCCATCCAATTCGATTCACCGCGGC 118588

QY 119 PheSerSerProThrGlnAspAlaMetValArgAspLeuAsnLeuSerIleSerGlu--- 137
Db 118589 TTCTCCTCACCCACCGAGCGATCGCCAGTCCCGACCTCAAGCTCTCCATCTCCGAGGTC 118648

QY 137 ----- 137

Db 118649 CCGTGCCCGTTCCAGCTGCGGATCTTGGCTGCTTTGCTGCTTCTCTGCTTATTGCGTT 118708

QY 138 -----PheSerAlaPheGlySerLeuSerAs 146
Db 118709 TCTTGTAGTGTGAAGTTTATCTTTCTGATGCGAGTTCTCGGCTTTCGGTTCGCTGCCAA 118768

QY 146 nValGlyGlyMetValGlyAlaIleAlaSerGlyGlnMetAlaGluTyrIleGlyArgLy 166
Db 118769 CGTCGGCGCCATGGTCGGAGCGATCGCCAGTGGCGAGATGGCGGAGTACATTGGCCGGAA 118828

QY 166 sGly----- 167

Db 118829 AGGGGTGAGAAATTTGTGCTTCATTGTTGGTTCCTTATCTGTTCTTCCACTAGAGGCCATCCT 118888

QY 167 ----- 167

Db 118889 GTATCTCGACGACGAACAAATTGATACTTTTGATAGTTGACAAGGTCAATACGTCACGTGTTG 118948

QY 167 ----- 167
Db 118949 TATGATTCCGGTAACACCAATAATTTGCCATTAAAGAAAGTGAGTGGGATTAGAAATGC 119008
QY 167 ----- 167
Db 119009 TGGTCATTTCCTGTGAACCTGTAGCTTGTGAAAAATATGAAACCCCAAAATTTTCAGGTTTCAGT 119068
QY 167 ----- 167
Db 119069 TAGCTATACTATTGGGGTTTGACCTTAATTTCAAAGTTGACGAGTCTTTTAGTTGTGATT 119128
QY 167 ----- 167
Db 119129 GCCATTGATCTGGCAAAAGAGGAGTTAGTATGACAGAAAATGTGTAGATGAAGTTTGT 119188
QY 167 ----- 167
Db 119189 GACATTTTCTTTGGTTTCATACTTTCATAATCACTTAATCAGATCAGTAGTTAACTACAGTG 119248
QY 168 -----SerLeuMetIleAl 172
Db 119249 AAGCTCTATTATTACTCTAAATGGTCTTGCTGATTGAATCTCCAGTCGTTGATAATTGC 119308

QY 172 aAlaIleProAsnIleIleGlyTrpLeuAlaIleSerPheAlaLys----- 187
Db 119309 GCGGTTTCCTAACATCATTTGTTGGCTTGCCATCTCCTTTGCAAAAGTACGTGTACACATA 119368
QY 187 ----- 187
Db 119369 TTTCTCAAGCTTATTTTCTCTGTTTGTCTCTTTTCACTAGTTTCACTCTTCCCGATCTT 119428
QY 188 -----AspAlaSerPheLeuTyrMetGlyArgLeuLeuGluGlyPheGlyValG1 204
Db 119429 TGTTTTTCAGGACGCGTCATTTTATACATGGGACGCTTGTCTGAAGTTTGGTGTGG 119488
QY 204 yIleIleSerTyrThr----- 209
Db 119489 TGTATATCATATACGCTCTGTATTTACCAACTTCATTTTCATGTTTATTGCTGCTTAGA 119548
QY 209 ----- 209
Db 119549 ATGTTCTGTGGCTTCACTGCTTGAGTACTTATCTGACTGAGACTGAGATGCGCTGTAGAAT 119608
QY 209 ----- 209
Db 119609 GAGGCAGAGCATCCCTTGCTAATGAAGCTTTTACATTTTGTGTCAGAGAAGAACTTTG 119668
QY 210 -----Va 210
Db 119669 GATCAGTTAATAGTTAGTGCATGTATAAACCCCTTAATGTGTGAATTATATCGTACCAGGT 119728
QY 210 lProValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerValas 230
Db 119729 GCCAGTATACATAGCAGAGATATCTCATCAGAACACAAAGAGGAGCACCTTGGCTCCGTGAA 119788

QY 230 nGln----- 231
Db 119789 CCAGGTGTGCTTATCGTCAGAAAGCATCTTATCTTATCGGAGGGAATTTATGGTTTGT 119848
QY 232 -----LeuSerValThrPheGlyIlePheL 240
Db 119849 ATGTTGGACTTACCATGCAGTTTTCATTATGTAGTTGTCCGTTACCATTTGGTATCTTGT 119908
QY 240 euAlaTyrLeuLeuGlyMetPheIleProTyrArgLeuLeuAlaValIle----- 256
Db 119909 TGGCCTATTGCTAGGCATGTTTGTCTTCTTGAGGCTGCTTGAGGTATAGGTACAAATC 119968
QY 256 ----- 256
Db 119969 AATACCTGATGCTGAATGTTTCTTATTAGGGATATGTAATAAATAAATGAATTTT 120028
QY 256 ----- 256

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis. 1 (bases 1 to 120977) Shinn,P., Brooks,S., Buehler,E., Chao,Q., Johnson-Hopson,C., Khan,S., Kim,C., Altafi,H., Bei,Q., Chin,C., Chiou,J., Choi,E., Conn,L., Conway,A., Gonzales,A., Hansen,N., Howing,B., Koo,T., Lam,B., Lee,J., Lenz,C., Li,J., Liu,A., Liu,K., Liu,S., Mukharsky,N., Nguyen,M., Palm,C., Pham,P., Sakano,H., Schwartz,J., Southwick,A., Thaveri,A., Toriumi,M., Vaysberg,M., Yu,G., Federspiel,N.A., Theologis,A. and Ecker,J.R. Genomic sequence for Arabidopsis thaliana BAC F18O14 from chromosome 1 Unpublished 2 (bases 1 to 120977) Ecker,J.R. Direct Submission Submitted (15-MAR-2000) Arabidopsis thaliana Genome Center, Department of Biology, University of Pennsylvania, 38th Street and Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA 3 (bases 1 to 120977) Ecker,J.R. Direct Submission Submitted (22-APR-2000) Arabidopsis thaliana Genome Center, Department of Biology, University of Pennsylvania, 38th Street and Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA 4 (bases 1 to 120977) Cheuk,R., Shinn,P., Brooks,S., Buehler,E., Chao,Q., Johnson-Hopson,C., Khan,S., Kim,C., Altafi,H., Bei,B., Chin,C., Chiou,J., Choi,E., Conn,L., Conway,A., Gonzales,A., Hansen,N., Howing,B., Koo,T., Lam,B., Lee,J., Lenz,C., Li,J., Liu,A., Liu,K., Liu,S., Mukharsky,N., Nguyen,M., Palm,C., Pham,P., Sakano,H., Schwartz,J., Southwick,A., Thaveri,A., Toriumi,M., Vaysberg,M., Yu,G., Davis,R., Federspiel,N., Theologis,A. and Ecker,J. Direct Submission Submitted (28-JUN-2000) Arabidopsis thaliana Genome Center, Department of Biology, University of Pennsylvania, 38th and Hamilton Walk, Philadelphia, PA 19104-6018, USA 5 (bases 1 to 120977) Chao,Q., Brooks,S., Buehler,E., Johnson-Hopson,C., Khan,S., Kim,C., Shinn,P., Altafi,H., Bei,B., Chin,C., Chiou,J., Choi,E., Conn,L., Conway,A., Gonzales,A., Hansen,N., Howing,B., Koo,T., Lam,B., Lee,J., Lenz,C., Li,J., Liu,A., Liu,J., Liu,S., Mukharsky,N., Nguyen,M., Palm,C., Pham,P., Sakano,H., Schwartz,J., Southwick,A., Thaveri,A., Toriumi,M., Vaysberg,M., Yu,G., Davis,R., Federspiel,N., Theologis,A. and Ecker,J. Direct Submission Submitted (11-OCT-2000) Arabidopsis thaliana Genome Center, Department of Biology, University of Pennsylvania, 38th and Hamilton Walk, Philadelphia, PA 19104-6018, USA On Apr 22, 2000 this sequence version replaced gi:7543634. Location/Qualifiers 1. 120977 /organism="Arabidopsis thaliana" /mol_type="genomic DNA" /db_xref="taxon:3702" /chromosome="1" /clone="F18O14" 90. 2159 /note="hypothetical protein" /codon_start=1 /evidence=not_experimental /product="F18O14.1" /protein_id="AAF79419.1" /db_xref="GI:8778411" /translation="MISFEVSPDVFTCSIWNAYCRSGNVDKAMVFAKETESSLGLLE NVVTYNSLINGAMIGDVEGMTRVLRMLSERGVSRNVVTYTSLIKGCKKGLMEAEH VFELKEXKLIVADQHMVGLMDGYCRTQIRDAVRVDNNMIEIGVRTNTTICNSLING YCKSGQLVEAEQIFSRMNDWSLKPDHHTYNTLVDGCRAGYVDEALKLCDQMCCKEV PTWMTYNTLLKGYSRIGAFHDVLSLWKVMKRGVNADEISCTSLLEALFKLGLDFNEAM KLWENVLARGLLTDTITLNVMSGLCKMEKVEAKEILDNVNIFRCKPAVQTYQALSH GYYKVGNLKAPAFVKEYMERKGIFFPTIEMYNLTLSGAFKYRHLNKVADLVIELRARGL TPTVATYGALITGWCNIGMIDKAYATCFEMIEKGITLNVNICSKIANSLFRLDKIDEA CLLLQKIVDFDLLPGYQSLKEFLEASATCLTKTQKIAESVENSTPKLLVPNNIVYN

VATGLCKAGKLEDARKLFSDLLSSDRFIPDEYTYTILIHGCAIAGDINKAFTLRDEM ALKGIIPNIVTYNALIKGLCKLGNVDRAQRLHLHLKLPQKGITPNAITYNTLIDGLVKSQ NVAEAMWLKEKMIKGLVRGSKQGDVDIPKEVVDPEVKLGSTGVIEMNSNELYDVR RVSEAVV" complement(4647. .5702) /note="putative glycosyltransferase-like protein emb|CAB42905.1; similar to EST gb|AI998490.1" /codon_start=1 /evidence=not_experimental /product="F18O14.2" /protein_id="AAF79456.1" /db_xref="GI:8778448" /translation="MSOHLILLILLLLHKKPISATTIIQKPKAAPQFYNsADCPLI DSESDDVVAKPIFCsRRRAVHVMTLDAAYIRGSVAAVLSVLQHSSCPENIVFHFVA SASADASSLRATISSsFPYLDFTVYFVFNSSVSRSLSSIRsSALDCPLNYARSYLADL LPpCVRVYLDSDLIILVDDIAKLAATDLGRDSVLAAPeYCNANFTSYFTSTFWSNPT LSLTFADRKACyENTGMVIDLSRWREGAYTSRIEeWMAMOKRMRIYELGSLPFPFLLV FAGLIKpVNHrWNQHGLGGDNFRGLCRDLHPGPVSLLHWSGKGKfPWARLDAGRPCPLD ALWAPYDLIQTpFALDS" complement(9620. .10300) /note="similar to Zn finger protein gi|6319911" /codon_start=1 /evidence=not_experimental /product="F18O14.3" /protein_id="AAF79455.1" /db_xref="GI:8778447" /translation="MSDVPsCSSGNDTNNDSSNFECNICLDLAQDPiVTLcGHLFCW PCLYKWLHLHSQSKDCpVCkAVIEEDRLVPLYGRKSSADPRsKSIpGLEVPNRPSGQ RPETAQDPDPNHGFAHHHGFGGMGfAAPMASARFGNVTLsAAFGGLIPSLFNHLHFH GFpDAAMYGAAsGGfPHGfSNPFHGSHMHsYQRHTGRGQGDHHLRILLIIVFVV VVFSLFLS" join(12523. .12577,12686. .13374) /note="hypothetical protein" /codon_start=1 /evidence=not_experimental /product="F18O14.4" /protein_id="AAF79420.1" /db_xref="GI:8778412" /translation="MAIFSTSHLLFIISFIATCTISVSgTTFLTnHCGSTIWPgILT ANGAQLGDGgPALASGSsVTFVSgWSGRFWARTYCNFDASGSgKCGTDCGSKLKC AGAGGAPPATLAeFTIGSGGKNAVQDFYDVSVLVDGYNVQMKITPQGSgDCKTAGCV SDVNAICPKELQVTGPsgVAAcKsACEAFNKPEYcCTGAYSTPATcPPTNYSKIFKQA CPSAYSAYDDASsTFTCTNANYEISfCS" complement(join(14020. .14136,14236. .14304,14402. .14464, 14557. .14696,14879. .15005,16274. .16486)) /note="unknown protein; similar to EST gb|H76493.1" /codon_start=1 /evidence=not_experimental /product="F18O14.5" /protein_id="AAF79454.1" /db_xref="GI:8778446" /translation="MLEAVDSsGVVNGGFpQIQsFYGDcSSEELsVLPRHTKVVVTG NNRTKsVLVLQGVVKAVGLGgWHVLVTNGIEVKLQrNALsVLEPPTGNEEDDDLD FENTQRNGSDMTsEDTLKPHKsKLRGQRsSRSSHKTMSRSLSSDSQsKSGFTPPENM KVDLSKLEMPALLNYWRHENLVDAIPNPSKEQLIDIVQRHFMSQQMDELQVIVGFVQA AKRMKCKACKFQsKESRNTDLNCIS" join(17994. .18489,18568. .18698,18814. .18847,19006. .19099, 19170. .19347,19441. .19479,19616. .19715,19855. .19994) /note="hypothetical protein" /codon_start=1 /evidence=not_experimental /product="F18O14.6" /protein_id="AAF79421.1" /db_xref="GI:8778413" /translation="MAKTDKLAQFLDsgIYESDEFNWFFLDTVrITNRSYTRFKVSPs AYsSRFFNSKQLNHsSESNPKKRKQKNsSFHLPSVGEQAsNLRHQEARLFLSKAH ESFLKEIELLsTKGLSDNDDDSSLNKcCDDDEVsFIELGgVQWQAFYEITLSFNL HCDNEGEScNQRFVQFVFNLVVNEIGEEVEAEfSNRRYIMPRNSCFYMSDLHHIRNL VPAKSEEGYNLIVIDPFWENASAHQKsFVIYLYMPTLPNQYFLSLPIKQLAHAEgAL VALWVTNREKLsFVEKELFPAWGIKYVATWYWLKVKPDGTLICDLDLLAGSEKRSDF KLDDKNQITINSIPGDFSRKPPIGDILLKHTPGSQPARCLELFAREMAAGWTsWGNEP L HFQDSRYFLKV" join(22037. .22300,22618. .23361)

/note="unknown protein; similar to ESTs gb|AI992723.1,
gb|T41777.1, dbj|AV440980.1, gb|AI997341.1, and
dbj|AV440785.1"
/codon_start=1
/evidence=not experimental
/product="F18014.7"
/protein_id="AAF79422.1"
/db_xref="GI:8778414"
/translation="MTSDGATSTSAIAAAMATRRKPSWRERENRRRRRAVAA
KIYTLRAQGNYNLPKHCNDNEVLKALCSEAGVVEEDGTYRKGHKPLPGDMAGSS
RATPYSSHNQSPLSPTSPILSYQVSPSSSPSPSRVGDPHNISTIFFLRNGGIP
SSLPLRISNSAPVTPVSSPTSRNPKPLPTWESFTKQSMMAKQSMTSLNYPFYAV
SAPASPTHRRQFHAPATIPEDCESDSTVDSGHWISFQKFAQQQPFASAMVPTSPFN
LVKPAQQLSPNTAAIQIGQSSEFKFENSQVWPWEGRIHVDVAMEDLELTGNGKAH
S"

CDS
complement(join(23868..24983,25237..25407))

/note="unknown protein; similar to EST gb|T13858.1,
emb|Z38050.1, and emb|Z38087.1"
/codon_start=1

/evidence=not experimental

/product="F18014.8"

/protein_id="AAF79453.1"

/db_xref="GI:8778445"

/translation="MAGRRDRSQQLRGSRIATAILIGIFIGVCVAVLPYPGFNSSSS
LKASEHLSKSSNQVSSACESPERVKMLKSDFTLSEKNAELKKQVRELTEKLRLAEQ
GSDNARKQVLALGTQIKAGPFGTVRSRNTPTILPDESINPLAKILEEIAVDKEIV
ALANANVKAMLEVOIASIKRVGITNYLVVALDDYIENLCKENDVAYYKRDPKDQDVT
GKTGNHAVSGLEKFRVLEFQLQGYGVLLSDVDIVFLQNPFSHLRDSDVESMSDGDH
NHTAYGNDVDEPAMGWARYAHTMRIWVNSGFFYLRTIPSIELLDRAVRLSKAK
VWDQAVNEELFYPSHPEYTHALHASKRVMDMYEFMNSKVLFTVRKNHELKKVKRPVI
VHVNYHPDKLNRMQAVVEFYVNGKQDALDSPDGSE"

CDS
complement(26100..27917)

/note="unknown protein; similar to ESTs gb|T42717.1,
emb|F13886.1, and emb|F13885.1"

/codon_start=1

/evidence=not experimental

/product="F18014.9"

/protein_id="AAF79452.1"

/db_xref="GI:8778444"

/translation="MADPLNGKSFIFCSILLFSFTLLFISPLVATESPVIEDVSTDVA
VSVSETNREAVLHNLEELVKNLTELVANLDAKLSATPLKEKNEISVDDDIIGEEKERG
RAKAFSVTKYSPFWSERFOFTSAVKLNSDATCINVLPRDFEGSSKYFAIGDSKGRVY
VFLRNGDVLIIEFTTVDSPVTAMVSYSVFKNSFVTVGHQNGAVLLHRIHEGNSGED
WNSNSVSMHVGVKFDVDSADPVTLLVEHVHVRVRYILATDLGKLTVLTENRTVYGS
VIPSSRPLVFLKQRLFLFTESGAGSLDRSMKIRETECEGLNHLNLTARTYVFDAAERSK

Alignment Scores:

Pred. No.: 3.64e-73 Length: 120977
Score: 1137.50 Matches: 335
Percent Similarity: 40.58% Conservative: 68
Best Local Similarity: 33.74% Mismatches: 78
Query Match: 40.93% Indels: 518
DB: 8 Gaps: 18

US-10-051-909-36 (1-553) x AC025808 (1-120977)

QY	52	GluGluSerGlySerAspHisAspGlyValLeuArgArgProLeuLeuAsnThrGlySer	71
Db	67805	GAAGAGGGAAGAATGAT-----CTTCGCGGACCGTTCCTTACACACAGGAAGT	67758
QY	72	TrpTyrArgMetSerSerArgGlnSerSerPheAlaProGlyThrSerSerMetAlaVal	91
Db	67757	TGGTATCGGATGGGTTTCGAGACAACTAGTATG-----TTGGAATCGTCTCAAGTT	67707
QY	92	LeuArgGluSerHisValSerAlaPheLeuCysThrLeuLeuValAlaLeuGlyProIle	111
Db	67706	ATTCGAGACAGCTCAATTTCTGTCTTAGCTTGTGTTTGTATGTTGCTCTTGGTCTATT	67647
QY	112	GlnPheGlyPheThrSer-----	117
Db	67646	CAATTGGATTCTACTGT-AAGAACACATAACAAATATCTCAACTTTCTTATATTTTT	67588
QY	117	-----	117

Db	67587	TTTATCAATGTGGATTCTCTGCTTTTGTCTCAAAATGACTAAGAAAGTGTGATTGTGAT	67528
QY	118	-----GlyPheSerSerProThrGlnAspAlaMetValArgAsp	130
Db	67527	TTATCTTTTGTGTTTTCAGTGTGGTTATCTCTCACCAACTCAAGCTGCCATTACTAAGGAT	67468
QY	131	LeuAsnLeuSerIleSerGlu-----	137
Db	67467	CTTGGTTAACTGTATCAGAGGTATGAAGTCCCATCTTTATTAAAGAACTGTTCTGTTT	67408
QY	137	-----	137
Db	67407	TTTTTTTTTTTTTTTGTAAAGACTCTTGTCTGTAAATTTCTGACTTTTAGGTTTCTTGT	67348
QY	138	-----PheSerAlaPheGlySerLeuSerAsnValGlyGlyMetValGlyAla	153
Db	67347	TTGATGTTTGTAGTACTCTGTGTTGGTCTCTATCCAAATGTGGTGCTATGTTGGAGCA	67288
QY	154	IleAlaSerGlyGlnMetAlaGluTyrIleGlyArg-LysGly-----	167
Db	67287	ATTGCGAGTGGTCAGATTGCTGAATACGTTGGACGGAAGGGGTATAATAGTTTGTGTTTC	67228
QY	167	-----	167
Db	67227	TTGATTTCTAAACAATGTTCTCATCTGTGACAAATTTGTCATGTAAACATCTTCTTCTTGT	67168
QY	168	-----SerLeuMetIleAlaAlaIlePr	175
Db	67167	GTAATCTTATTACTGTGTTTCTTCTTCTTGTGACAGTCTCTGATGATTGTCGAATTC	67108
QY	175	oAsnIleIleGlyTyrLeuAlaIleSerPheAla-Lys-----	187
Db	67107	CAATATTATTGGATGGCTTTCGATATCATTTGCAAAAGTGAGTGTGTTTGGTGTCTCAAA	67048
QY	187	-----	187
Db	67047	GTTTCTGATATTCTTGGCTTATCTCTATACCTACTAAGTTGTTGTTGTTCTTGGTTTGT	66988
QY	188	-----AspAlaSerPheLeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleI	206
Db	66987	TTTAGGATACTTCTTTCTTACATGGGAAGATTGTAGAAGCTTCGGAGTTGGGATTA	66928
QY	206	LeSerTyrThrVal-----	210
Db	66927	TCCTCTATACGGT-AATATTAGATTCAACTATTTCTTGTGTTTCAGCAATGTTGTTGG	66869
QY	211	-----ProValTyrIleAlaG	216
Db	66868	AGTGATCAATATAATATACACACCTTTTGAATTGCGATTTTAGGTTCCCGTATATAGCTG	66809
QY	216	IuIleSerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeu-----	232
Db	66808	AGATCGCTCCACAGACCATGAGAGGAGCCCTTAGGTTAGTTAACCCAGGT-TTGTCAATGT	66750
QY	232	-----	232
Db	66749	TAGTTAGCTTCACTATTTTATGTTTTCATTTTATGAGCTGATTTCATATGTTAACTAAA	66690
QY	233	-----SerValThrPheGlyIlePheLeuAlaTyrLeuL	244
Db	66689	CAATAAATGTTTGTGTTTATAGCTTCTCTGTAAACAATTGGGATAATGCTGCGGTATTAC	66630
QY	244	euGlyMetPheIleProTyrArgLeuLeuAlaVal-----	255
Db	66629	TCGGTCTCTTTGTTCCATGGAGAATTTCTTGCAGTTTGGGTAACTCACAACAACTTTTGA	66570
QY	255	-----	255
Db	66569	CTTGACTTGGAGAACTATACCTGTTTTTATTTTATTTTATTTGATCGATACAGTCTAATCAC	66510
QY	256	-----IleGlyAlaLeuProCysThrMetLeuIleProGlyL	268

Db 66509 TTTCTTGTGTCCTCTTTATGTTTAGGAGTATTGCCAAGTACATGTTGATACCGGGTC 66450

Qy 268 euphepelleProGluSerProArgTrpLeu----- 278

Db 66449 TTTTTCATTCCTGAATCTCCTCGGTGGTGGTATGCCCTCFACAACCGTATTGCTTATA 66390

Qy 279 -----Ala 279

Db 66389 TTGTTGAGTTGTGTCATTATTGGTTATCCTAAATTTGTTGGTGTGTAATTGTAGGCA 66330

Qy 280 LysMetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPheGluThr 299

Db 66329 AAGATGGGTTGACAGATGATTTCGAAACTTCATTGCAAGTTCCTCGTGGATTGAGACT 66270

Qy 300 AspIleThrThrGluValAsnAspIleLys----- 309

Db 66269 GATATTACCGTTGAGGTTAATGAATCAAGGTAGTGACTAAACTGAAGAAATGCTTTGAT 66210

Qy 310 -----ArgAlaValAlase 314

Db 66209 GTGATTTCTTGTCTTATTATAAATTTTCATTTTATGCACCTCTGCAGAGATCTGTGGCATC 66150

Qy 314 rSerSerLysArgThrThrIleSerPheGlnGluLeuAsnGlnLysLysTyArgThrPr 334

Db 66149 ATCTAGCAACGTTCTGCAGTTCGGTTTGTAGACCTCAAGCGCAGGAGATACTATTTCCTCC 66090

Qy 334 oLeuLeu----- 336

Db 66089 ACTGATGGTGCAATTTCTCTTACCTCTCTTGAACATAATTTTGTCAACACATGAGAGAAAC 66030

Qy 337 -----LeuGlyIleGlyLeuLeuValLeuGlnAsnL 347

Db 66029 TCACTCTTTTCTGTAACTCTGTGTTACAGGTTGGTATAGGGCTGCTGCACCTCAACAAC 65970

Qy 347 euSerGlyIleAsnGlyValLeuPheTyAlaSerSerIlePheLysAlaAla----- 364

Db 65969 TTGGAGGAATCAATGGTGTCTTGTCTATTCGAGTACAATTTTGAATCTGC-AGGTAGA 65911

Qy 364 ----- 364

Db 65910 GTCATATGTTTTTTTGTGATTATTGGAATGGTCCTAAGATAGATAAACAAGATTTCTAACT 65851

Qy 365 -----GlyValThrAsnSerAspLeuAlaThrC 374

Db 65850 ATGCTTCTTTGTGCAATTATCATTCACCTAAGGGGTGACATCAAGTAATGTGGCGACAT 65791

Qy 374 ysSerLeuGlyAlaIleGlnValLeu----- 382

Db 65790 TTGGAGTTGGCGTGTTCAGGTATA-TCATTTCAATCTAGGAGACATTAGATTCGATTTT 65732

Qy 383 -----AlaThrGlyValThrTrpLeuLeuA 392

Db 65731 TCTAACCGTTGTTCATCTATACCAGGTAGTGGCGACTGGGATAGCGACATGGTTGGTGG 65672

Qy 392 spArgAlaGlyArgArgIleLeuLeuIle----- 401

Db 65671 ATAAAGCAGGTCGTGGCTTCTACTCATGGTGAAGAAATATTTGAGATTCATCTGTATATC 65612

Qy 401 ----- 401

Db 65611 TTACAAAGCTGAAGAATTTATGTATGCATAACATAATATTAGCATTTTCTTCTGAAATG 65552

Qy 402 ---IleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSerValValPhePheLeu 420

Db 65551 CAGATCTCTTCTATCGGAATGACGATAAGCCTCGTGTATTGTAGCATTTTACCTT 65492

Qy 421 Lys----- 421

Db 65491 AAGGTAACATTTTAACATTTTGGACATCTCTACTTGTACAAATGTAAAGTGTCACTCTAATT 65432

Qy 422 -----AspAsnIleSerGlnAspSerAsnSer 430

Db 65431 ATGAGATATTGAGGTTCTTAATCTTTTGTAGGAATTTGTATACCTGATTCCAACATG 65372

Qy 431 TyrTyrIleLeuThrMetIleSerLeuValGlyIleVal----- 443

Db 65371 TACAACATTTCTAAGCATGGTTTCCGTAGTTGGAGTTGGTAAGTCGTTGTTTATACA 65312

Qy 443 ----- 443

Db 65311 TTCACACAACGAAAGATACTTTCTTATTATTACATATTTCCCGGTATGTGTAATAATCTGG 65252

Qy 444 -----SerPheValIleThrPheSerPheGlyMetGlyAlaIleProTrp 458

Db 65251 CTGAGTTATTACAGGCTATGGTTATTTCTTCCTCTCTAGGAATGGGACCAATTCATGG 65192

Qy 459 LeuMetMetSerGlu----- 463

Db 65191 CTGATTATGCTGAGGTACTAAATATTATAAACATGGTGTTCCTTAGTTATAAATCACA 65132

Qy 464 -----IleLeuProValSerIleLysSe 471

Db 65131 CTGGAATAATTTAATACTGATTAATGTTGGTAAACAGATTTCTGCCAGTGAATATAAGGG 65072

Qy 471 rLeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeuThrSerPheAlaIleThrMetTh 491

Db 65071 TTTAGCCGGAAGTATAGCAACTTTGTTAAACTGGTTCGTGTCATGGTTAGTGACAATGAC 65012

Qy 491 rThrAsnLeuMetLeuThrTrpSerValGlyGlyThrPheLeuSerTyrMetValValSe 511

Db 65011 TGCAAAATATGCTGTAGCATGGAGCAGTGGGGAACTTTCACTCTCTACGCTTTGGTTTG 64952

Qy 511 rAlaPheThrIleValPheValValLeuTrpValProGluThrLysGly***AsnSerAr 531

Db 64951 TGGATTCACTGTGGTTTGTGAGTCTTTGGGTCTCTGAGACTAAAGGAAA-AACGCTTG 64893

Qy 531 gGlyAspThrIlePheValSerLeu 539

Db 64892 AAGAGATCCAAGCTTTGTTTCAGATG 64868

RESULT 15

AC025814

LOCUS

DEFINITION Arabidopsis thaliana chromosome 1 BAC F22H5 genomic sequence, complete sequence.

ACCESSION AC025814

VERSION AC025814.7 GI:12331602

KEYWORDS HTG.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

REFERENCE 1 (bases 1 to 68041)

AUTHORS Lin,X., Kaul,S., Town,C.D., Benito,M.-I., Creasy,T.H., Haas,B.J., Wu,D., Maiti,R., Ronning,C.M., Koo,H., Fujii,C.Y., Utterback,T.R., Barnstead,M.E., Bowman,C.L., White,O., Nierman,W.C. and Fraser,C.M.

TITLE Arabidopsis thaliana chromosome 1 BAC F22H5 genomic sequence

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 68041)

AUTHORS Lin,X. and Kaul,S.

TITLE Direct Submission

JOURNAL Submitted (15-MAR-2000) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA, xlin@tigr.org

REFERENCE 3 (bases 1 to 68041)

AUTHORS Town,C.D. and Kaul,S.

TITLE Direct Submission

JOURNAL Submitted (12-SEP-2000) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA, cdtown@tigr.org

REFERENCE 4 (bases 1 to 68041)

AUTHORS Town,C.D. and Kaul,S.

TITLE Direct Submission

JOURNAL Submitted (05-OCT-2000) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA, cdtown@tigr.org

REFERENCE 5 (bases 1 to 68041)

AUTHORS Town,C.D. and Kaul,S.

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Direct Submission
Submitted (12-OCT-2000) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA, cdtown@tigr.org
6 (bases 1 to 68041)
Town, C.D. PhD.
Direct Submission
Submitted (22-JAN-2001)
On Jan 22, 2001 this sequence version replaced gi:12280794.
Address all correspondence to: at@tigr.org

BAC clone F22H5 is from Arabidopsis thaliana chromosome 1
The orientation of the sequence is from SP6 to T7 end of the BAC
clone.
Genes were identified by a combination of several methods: Gene
prediction programs including Genscan+ (Chris Burge,
http://CCR-081.mit.edu/GENSCAN.html), GeneMarkHM (Mark Borodovsky,
http://genemark.biology.gatech.edu/GeneMark/), GlimmerA (a variant
of GlimmerM, see Mihaela Pertea,
http://www.tigr.org/softlab/glimmerm.htm/glimmerm.html, and
GeneSplicer (Mihaela Pertea and Steven Salzberg, contact
mpertea@tigr.org), searches of the complete sequence against a
peptide database and the plant EST database at TIGR
(http://www.tigr.org/tcb/tgi.shtml). Annotated genes are named to
indicate the level of evidence for their annotation. Genes with
similarity to other proteins are named after the database hits.
Genes without significant peptide similarity but with EST
similarity are named as unknown proteins. Genes without protein
or EST similarity, that are predicted by more than two gene
prediction programs over most of their length are annotated as
hypothetical proteins. Genes encoding tRNAs are predicted by
tRNAscan-SE (Sean Eddy, http://genome.wustl.edu/eddy/tRNAscan-SE/).
Simple repeats are identified by repeatmasker (Arian Smit,
http://ftp.genome.washington.edu/RM/RepeatMasker.html).

Location/Qualifiers
1. .68041
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/cultivar="Columbia"
/db_xref="taxon:3702"
/chromosome="1"
/clone="F22H5"
complement(418..5404)
/gene="F22H5.15"
complement(join(<418..594,684..773,873..938,1034..1095,
1179..1227,1319..4820,5040..5404))
/gene="F22H5.15"
complement(join(<418..594,684..773,873..938,1034..1095,
1179..1227,1319..4820,5040..5404))
/gene="F22H5.15"
/codon_start=1
/product="unknown protein, 3' partial; 5120-134"
/protein_id="AAG12690.1"
/db_xref="GI:10092277"
/translation="WEYEKSPATTTFSRKISNNRSHLSFSANAVYDGVFSSPVNSKS
PLVDYGEIFRGSPSPSSIPFLDVPENLVGVKVDVRSKLDYSSVFGGLGACDFAVT
PKEVIIKSEKTSINEDKRRNRKGGSSDPLCNEGKRSPEVMVMKHSDDISYHQTV
RNENGATHLTQVATWPGPIPTQVVDNTSLHKIESKSTPIPAVEKKLPCENEGREEVK
ASRKQSGKTEVDENIFARDGCTRDDSTCKTVSNGEYRDVKPPSSFQCTLNGEHGAS
ERLGLNSGSPSERYETEDADSPSPPYFAETDENSVAEASSALKKAEAEQIRMNI
AKOMMEKKSGFRSCAKLKSCDDSKIENKNTKVEGITEESRDNNNSQILGEMVKPSEQ
SFSNEGDAQHAKARKLWGVPEGLLKSTSDHKPEELEEQDIITLBEAQARRGRKHVELP
GGIFKSVMSKQOEPENLAPAKPEPDKQVQPIITENPFYFGLGSKLKCVEAFSTG
SKVSQKDEKQFTENKSTVTQMVEDSSQEMLAGIPVIETYLREVEETPQOTES
EMNIEEKSESTICAPTERSOMMEKETGWKVSACKFEDGSGVKDFQENGDTCTNVL
QEGEKEIVSEPOMLVGPDDSKTYVREVEETPTPSLNKTQSDDSVGAUVSNRNVISE
PGNIDVEQVAVHKVPRRRRWKTSQEDVVMKAPKGNRPWOLFSAENETTAMSFHEE
GVRIHHAASEIESTSQASDGLQENWTVLKQFRMQFQATDCKEDTETCLVESERG
HLDIHQAKQEQVEQVETVRTNYEAYATRENEDESAQETCYCKBDGKVEVQGTSL
VRELIQEELMASLEEDVQEAEEAGWVQGLSELNEINEHADSHAEMLLEYDRSETDS
NNSRERFDQTQEQAEETMIDGSDITDTSRSSPFMRQGDYSIEEVTEIQRHRSQDFPEKA
SAVSNTTEHIEEIDSISIQSGWSVVEDDDRSQDGGASQAEKHELEETKEESDEMK
TSLGVERNGDKKLELHQFEQENETYSRVNVEAAESSCRFPNGEETIGAAATNGMKENE
GEEGEESCRSSMEEGDATSDIGAATDGNMKNENEGEESCRSSMEEGDATSDISQNK

FEATURES
source

gene

mRNA

CDS

ETVEHLKKIDETREKERERKQERVVVERAIRERARAFADAMERACKTAMEKAKAVA
HREVPKSEKSGVEVNDKLSSAEKASMOAKLRAERAAVERAITEVREERAMEKALSGK
SAASKSYGGKSPSSGERRGSSGTEKSSGSPSNSSNQTEGFIQRCRKARSERHQ
RTSDRAAEALAEKKLRDLTKQEQTERNLAEALDADVKRWSSGKENLRALISTLOY
ILGAESGWKPIPLTDLVSSASVRKAYRKATLYVHPDKLQQRGASTQQKYICEKVFDLL
K"
complement(9354..10899)
/gene="F22H5.16"
/note="similar to GI:1708420 from [Arabidopsis thaliana]
(J. Biol. Chem. 270 (44), 26224-26231 (1995))"
complement(join(<9354..9638,9750..9953,10043..10328,
10461..10595,10671..>10899))
/gene="F22H5.16"
complement(join(9481..9638,9750..9953,10043..10328,
10461..10595,10671..10856))
/gene="F22H5.16"
/codon_start=1
/product="NADPH oxidoreductase, putative; 10572-9197"
/protein_id="AAG12695.1"
/db_xref="GI:10092282"
/translation="MTTEKSKILVIGGTGYMGFEFIVEGSAKAGNTPFALVREASLSDP
VKSRTIQSFKDLGVTLHGLDNHESLVKAIKQVDVISTIGHKQIFDQTKIISAIKE
AGNVKRFLEPAEFGIDVERTSAVEPAKSLFAGKVQIRRAIEAGIPYTVVSNCSAGFY
LRTLLQFESGLSHTRDKAIIFGDKNVPPRDVKVTLIDGNKAVVINKEEDVAAAYMIKA
VDDLRLNKLTYISPPNNILSMNEMVTLWEKKIGKLSLEKTHISEEQLKSIQVPIDVF
KSINHAVFVKGDQTSFTIEPWFGEESVLYPDVKYTSIDEYLSQFT"
complement(11122..12573)
/gene="F22H5.18"
/note="similar to GI:1708420 from [Arabidopsis thaliana]
(J. Biol. Chem. 270 (44), 26224-26231 (1995))"
complement(join(<11122..11422,11516..11734,11791..12034,
12117..12251,12333..>12573))
/gene="F22H5.18"
complement(join(11235..11422,11516..11734,11791..12034,
12117..12251,12333..12518))
/gene="F22H5.18"
/codon_start=1
/product="NADPH oxidoreductase, putative; 12234-10951"
/protein_id="AAG12680.1"
/db_xref="GI:10092267"
/translation="MASEKSKILVIGGTGHIGKLIIEASVKAGHSTLALVREASLSDP
NKGKTQVNFKDFGVTLHGLDNHESLVKAIKQADVISTVSGMQILDQTKIISAIKE
AGNVKRFLEPSEFGMDVDKSSAVEPAKSLFAGKQLQRRDIEAEGIPYTLVNTYFAGYY
LPTLVQLEPGLTSPPRDKVKIFGDKNVKVEYFIAVINKEEDIAAYTIKAVDDPRTLNK
TLIYNPPNNTLSMNEIVTLWEKKIGKSVKIMSEEQIFKSIQESPVFFNVLLSINHA
VFVKGDQTNFTIEPSEFGEASELYPDIKYTSIDEYLSYFALGTSLNT"
complement(12872..14426)
/gene="F22H5.17"
/note="similar to GI:1708420 from [Arabidopsis thaliana]
(J. Biol. Chem. 270 (44), 26224-26231 (1995))"
complement(join(<12872..13216,13298..13513,13578..13821,
13974..14108,14193..>14426))
/gene="F22H5.17"
complement(join(13053..13216,13298..13513,13578..13821,
13974..14108,14193..14378))
/gene="F22H5.17"
/codon_start=1
/product="NADPH oxidoreductase, putative; 14094-12769"
/protein_id="AAG12677.1"
/db_xref="GI:10092264"
/translation="MATEKSKILVIGGTGYIGKFLVEASAKAGHSTPALVREATLSDP
VKGTVQSFKDLGVTLHGLDNHESLVKAIKQVDVISTVSGMQILDQTKIISAIKE
AGNVKRFLEPSEFGVDVDRTSAVEPAKSLFAGKIQIRRTIEAEGIPYTAVTGCFGGYY
LPTLVQFEPGLTSPPRDKVTILDGNAKBYFAAVINKEEDIAAYTIKAVDDPRTLKI
LYIKPSNNTLSMNEIVTLWEKKIGKSLKTHLPEEQLLKSIQESPIPNVLSINHA
FVNGDTNISIEPSFGEASELYPDVKYTSVDEYLSYFA"
complement(16153
15081..16153)
/gene="F22H5.1"
/note="similar to GI:6939839 from [Oryza sativa]"
join(15081..15228,15380..15470,15661..>16153)
/gene="F22H5.1"
join(15171..15228,15380..15470,15661..16153)
/gene="F22H5.1"

gene

mRNA

CDS

```
/codon_start=1
/product="GSH-dependent dehydroascorbate reductase 1,
putative; 14887-15869"
/protein_id="AAG12679.1"
/db_xref="GI:10092266"
/translation="MALDICVKVAVGAPDVLGDCPFQSRVLLTLEKKLPYKTHLINV
SDKPQWFLDISPEGKVPVVKLDKQWVADSDVIGLLEKVPPEPSLKTPEFASVGSKI
FGAFVTLKSDANDGSEKALVDELEALENLKTHSGPFVAGEKITAVDLSLAPKLYH
LEVALGHYKWNVSPELSLTVRNKALFSRESFENTKAKKEIVVAGWESKVNA"
complement(16410..18554)
/gene="P22H5.2"
complement(join(<16410..16558,16703..16830,17050..>18554))
/gene="P22H5.2"
complement(join(16410..16558,16703..16830,17050..18554))
/gene="P22H5.2"
/codon_start=1
/product="unknown protein; 18270-16126"
/protein_id="AAG12682.1"
/db_xref="GI:10092269"
/translation="MP2FRFRIPFFSSNSPSRISSTSSPSPPTTPPSRPPFRPAGIA
```

Alignment Scores:				
Pred. No.:	5.15e-72	Length:	68041	
Score:	1117.50	Matches:	340	
Percent Similarity:	38.10%	Conservative:	73	
Best Local Similarity:	31.37%	Mismatches:	82	
Query Match:	40.21%	Indels:	595	
DB:	8	Gaps:	18	
US-10-051-909-36 (1-553) x AC025814 (1-68041)				
QY	52	GluGluSerGlySerAspHisAspGlyValLeuArgArgProLeuLeuAsnThrGlySer	71	
Db	33823	GAGGAGGCTAGGAATGAT-----TTGCGTCGGCCGTTTATTACACCCGGAAGC	33870	
QY	72	TrpTyrArgMetSerSerArgGlnSerSerPheAlaProGlyThrSerSerMetAlaVal	91	
Db	33871	TGGTACCGGATGGGTCGAGACAATCGAGTATG-----ATGGGTTCTCTCAAGTT	33921	
QY	92	LeuArgGluSerHisValSerAlaPheLeuCysThrLeuIleValAlaLeuGlyProIle	111	
Db	33922	ATTCGAGACAGTCTATCTCTGTTCTTGCAATGTGTTTGTGATCGTTCCTGGCCCTATT	33981	
QY	112	GlnPheGlyPheThr-----	116	
Db	33982	CAATTCGGATTCACTGTGAAGACAAATTCAAAATCTCAACTTTTTTTTACATTTTATTCCC	34041	
QY	116	-----	116	
Db	34042	TTGTAGAAGAAAGGTGTGACTTTTGGTGGAAAAAAGTAGTGTAGCTGGATTGATGATC	34101	
QY	116	-----	116	
Db	34102	TTGGGTTTGGTTTGATTGTGATTTTCAATGCAAGTCTGGTGATTGAAAAATCTCATATT	34161	
QY	116	-----	116	
Db	34162	TGGTGAACCGAGATGATTTTTCAGGTTTGCTGACAGATTTAGAGGACTTTGAATTGGGA	34221	
QY	116	-----	116	
Db	34222	ATTTCGATTGTTTGCTTAAACCTTCCATGTTTGACTTTTGATCACCATCTTTTAGTACTC	34281	
QY	117	-----	124	
Db	34282	TGTGTTTGGTCCTTAAACCCCTCCATCCTTTTGTTCAGTGTGGTTATTATCATCTCCAAC	34341	
QY	124	nAspAlaMetValArgAspLeuAsnLeuSerIleSerGlu-----	137	
Db	34342	AGCTGCAATCACCAGGATCTCGGTTTAAACAGTATCTGAGGTAATGAGCATCATTTGTC	34401	
QY	137	-----	137	

Db	34402	TAGAGCTGTCTGGTTTTTATCATTTGCTTGTGTTTAGCCGTTGTTTGATGTACCACTC	34461	
QY	138	-----PheSerAlaPheGlySerLeuSerAsnValGlyGlyMetValGlyAlaIle	154	
Db	34462	CATGTTTTAGTACTGTGTTTGGTTCTCTTTCCAATGTGGGTGCTATGTTGGTGCAAT	34521	
QY	154	eAlaSerGlyGlnMetAlaGluTyrIleGlyArg-LysGly-----	167	
Db	34522	TGCCAGTGGTCAGATTGCAGAAATACATTGGACGAAAAAGGGGTATACTTTGTTATCTGATT	34581	
QY	167	-----	167	
Db	34582	CTTGATCGTTCTGTTTATAGATCCATTTTATAAAGTCTCTTTCTTGGGTTCTAAACATCT	34641	
QY	168	-----SerLeuMetIleAlaAlaIleProAsnIleIleGlyTyrL	181	
Db	34642	TATTGTGTTTTCTTGGACAGTCTCTTATGATTGCTGCAATTCCCAATATATCATTTGGATGGC	34701	
QY	181	euAlaIleSerPheAlaLys-----	187	
Db	34702	TTTGCAATATCATTTGC-AAAAGTGAGTTTTTGTGCTTCTCATACCATAAATTGGCGATT	34760	
QY	188	-----AspAlaSerPheL	192	
Db	34761	GATTGTGTTGCTTGACTTTAACTCTTTTACATCATTTCTGGTTTGTAGGATACTTCTTTTC	34820	
QY	192	euTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTyrThr-----	209	
Db	34821	TTTACATGGGAAGACTTTTAGAAGGCTTTGGCGTTGGGATCATCTCTTACACTGTAATAC	34880	
QY	209	-----	209	
Db	34881	TGAACCTTAAGTATTTACTTTATGGAATAACGGTTCAGTATATTAATGCAACTCGTT	34940	
QY	210	-----ValProValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaL	226	
Db	34941	CGTGATTTTAGGTACTGTATATATCGCTGAGATAGCTCCACAAAATATGAGAGGAGGAT	35000	
QY	226	euGlySerValAsnGlnLeu-----	232	
Db	35001	TGGGTTTCGGTTAACCAAGGT-TAGACATGTGTAGTACTAATTCCTGTTTTCAGCTTTTGTG	35059	
QY	233	-----SerValT	235	
Db	35060	TACTGACTCGTAATGCTAAACAAGAAATGTTGTGTAACCTTCTTGTGCTGAGCTTTCTGTGA	35119	
QY	235	hrPheGlyIlePheLeuAlaTyrLeuLeuGlyMetPheIleProTyrArgLeuLeuAlaV	255	
Db	35120	CAATTGGAATAATGTTGGCGTATTTACTTGGCCTCTTTGTTCCATGGAGAAATCTTGCAG	35179	
QY	255	al-----	255	
Db	35180	TTCTGGGTAAATCTCATTTACACGGCCTGAGACTATTTTGTGATTATGAATAATTGCTAC	35239	
QY	256	-----IleGlyAlaLeuProCysThrMetL	264	
Db	35240	GCTCTAATCACITTTCTGTTCTCTTCTCTTTTATCTTTAGGGATATTGCCCGTGTACATTAT	35299	
QY	264	euIleProGlyLeuPhePheIleProGluSerProArgTyrLeu-----	278	
Db	35300	TGATACCAGGTCTCTTTTTCATCCCCGAATCCCTCGCTGGTGGTAAGCCTTACAGTAG	35359	
QY	278	-----	278	
Db	35360	TTTGCTTATTCTTCTTGAGTTTAGTGCAATATTTGGTTACTCCACTCTTTTTTTTTTTTT	35419	
QY	279	-----AlaLysM	281	
Db	35420	TAATAACATCATTAATCATCTCTGTATGGGATACTTGTATTTCTGTTAAATGTAGCAAAAA	35479	
QY	281	etAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPheGluThrAspI	301	
Db	35480	TGGGTATGACAGATGAATTGAAACTTCATTACAAAGTTCTTCGAGGATTTTGAGACTGATA	35539	

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
OM protein - nucleic search, using frame_plus_p2n model
Run on: October 13, 2004, 11:39:19 ; Search time 676.434 Seconds
(without alignments)
4291.521 Million cell updates/sec

Title: US-10-051-909-36
Perfect score: 2779
Sequence: 1 PSSSSFRPAGKKKKKQNG.....TIFVSLSIQRQLWLPCLSL 553

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-O=/cgn2 1/USPTO spool.h/US10051909/runat 13102004 123335 19900/app_query.fasta_1.1678
-DB=N_Geneseq_23Sep04 -QFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10051909@CGN_1.1.891@runat 13102004 123335 19900 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :				N_Geneseq_23Sep04:*			
				1:	geneseqn1980s:*		
				2:	geneseqn1990s:*		
				3:	geneseqn2000s:*		
				4:	geneseqn2001as:*		
				5:	geneseqn2001bs:*		
				6:	geneseqn2002as:*		
				7:	geneseqn2002bs:*		
				8:	geneseqn2003as:*		
				9:	geneseqn2003bs:*		
				10:	geneseqn2003cs:*		
				11:	geneseqn2003ds:*		
				12:	geneseqn2004s:*		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

				SUMMARIES			
Result No.	Score	Query Match	Length DB ID	Description			

1	2777	99.9	1953 12	ADG47939	Adg47939 Corn Beta		
2	1710	61.5	1377 8	ADA69666	Ada69666 Rice gene		
3	1577	56.7	1428 8	ADA69667	Ada69667 Rice gene		
4	951.5	34.2	1741 3	AAC36319	Aac36319 Arabidops		
5	943.5	34.0	1683 3	AAC39099	Aac39099 Arabidops		
6	933	33.6	1395 3	AAC45857	Aac45857 Arabidops		

RESULT 1	
ADG47939	
ID	ADG47939 standard; cDNA; 1953 BP.
XX	AC ADG47939;
XX	AC ADG47939;
DT	11-MAR-2004 (first entry)
XX	Corn Beta-vulgaris-like sugar transport protein cDNA #4.
DE	Arabidopsis-like sugar transport protein;
XX	Beta-vulgaris-like sugar transport protein; transgenic; physical mapping;
KW	corn; plant; gene; ss.
KW	
KW	
XX	
OS	Zea mays.
XX	
Key	Location/Qualifiers
PH	1..1662
FT	/*tag= a
FT	/product= "Corn Beta-vulgaris-like sugar transport
FT	protein"
FT	/note= "Xaa can be any amino acid, CDS does not include
FT	no start codon"
FT	/transl except= (pos:1582..1584, aa:Xaa)
FT	/partial
FT	
XX	
PN	US2002199217-A1.
XX	
XX	
PD	26-DEC-2002.
XX	
XX	

7	923	33.2	1685	3	AAC45298	Aac45298 Arabidops
8	923	33.2	1743	3	AAC45290	Aac45290 Arabidops
9	920.5	33.1	1473	6	ABZ14526	Abz14526 Arabidops
10	916	33.0	1755	3	AAC40459	Aac40459 Arabidops
11	899	32.3	1359	6	ABZ13822	Abz13822 Arabidops
12	893	32.1	1768	12	ADMA7939	Adm47939 Polynucle
13	877.5	31.6	1407	6	ABZ13486	Abz13486 Arabidops
14	877.5	31.6	1407	8	ADA67841	Ada67841 Arabidops
15	877.5	31.6	1642	3	AAC46009	Aac46009 Arabidops
16	868.5	31.3	1425	3	AAC50050	Aac50050 Arabidops
17	853.5	30.7	1425	3	AAC50049	Aac50049 Arabidops
18	837.5	30.1	1209	8	ADA70113	Ada70113 Rice gene
19	824	29.7	1452	3	AAC42659	Aac42659 Arabidops
20	805.5	29.0	1695	3	AAC49517	Aac49517 Arabidops
21	790.5	28.4	1436	3	AAC45400	Aac45400 Arabidops
22	786.5	28.3	1440	3	AAC33666	Aac33666 Arabidops
23	777	28.0	1374	6	ABZ13215	Abz13215 Arabidops
24	758.5	27.3	1378	3	AAC35123	Aac35123 Arabidops
25	754.5	27.2	1344	8	ADA69777	Ada69777 Rice gene
26	688.5	24.8	1724	4	ABL13389	Ab113389 Drosophil
27	682.5	24.6	1239	3	AAC34742	Aac34742 Arabidops
28	674.5	24.3	2072	4	AAF55867	Aaf55867 Murine GL
29	674	24.3	2217	4	AAF55865	Aaf55865 Human GLU
30	672	24.2	898	10	ADD17645	Add17645 DNA (Seq1
31	672	24.2	2087	4	AAF55866	Aaf55866 Rat GLUTX
32	669.5	24.1	2080	4	AAD09552	Aad09552 Human tra
33	663	23.9	1473	6	ABZ13864	Abz13864 Arabidops
34	663	23.9	1473	6	ADG88039	Adg88039 A. thalia
35	658.5	23.7	1445	8	ABZ24794	Abz24794 Human sol
36	658.5	23.7	1856	8	ABZ24792	Abz24792 Human sol
37	652.5	23.5	1873	8	ABZ24793	Abz24793 Human sol
38	652.5	23.5	2514	4	ABL17067	Ab117067 Drosophil
39	640	23.0	1412	4	ABL13883	Ab113883 Drosophil
40	623.5	22.4	2487	8	ABZ24790	Abz24790 Human sol
41	621.5	22.4	1541	4	AAF55870	Aaf55870 Human GLU
42	621	22.3	2356	8	ADA84076	Ada84076 Human SLC
43	621	22.3	2356	8	ABZ24791	Abz24791 Human sol
44	621	22.3	2390	4	AAK94338	Aak94338 Human ful
45	621	22.3	2390	12	ADL31000	Adl31000 Full leng

Db	1159	TCCTTCGGCATGGTGCCATTCCATGGGTCA	TGATGCTGAGATCCTCCACGTTAGCATC	1218	PR	25-MAY-1999;	99US-0136021P.
					PR	27-MAY-1999;	99US-0136392P.
					PR	28-MAY-1999;	99US-0136782P.
Qy	470	LysSerLeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeuThrSerPheAlaIleThr		489	PR	01-JUN-1999;	99US-0137222P.
					PR	03-JUN-1999;	99US-0137528P.
Db	1219	AAGAGTCTGGGGGAAGCTTTGGCAGACTTGCCAAACATGCTTACATCCTGGGCAATACA		1278	PR	04-JUN-1999;	99US-0137502P.
					PR	07-JUN-1999;	99US-0137724P.
Qy	490	MetThrThrAsnLeuMetLeuThrTrpSerValGlyGlyThrPheLeuSerTyrMetVal		509	PR	08-JUN-1999;	99US-0138094P.
					PR	10-JUN-1999;	99US-0138540P.
Db	1279	ATGACAGCAAAATTTGTTACTCAGCTGGAGTCTGGAGGAACATTTTGTCTCTACATGATT		1338	PR	10-JUN-1999;	99US-0138847P.
					PR	14-JUN-1999;	99US-0139119P.
Qy	510	ValSerAlaPheThrIleValPheValValLeuTrpValProGluThrLysGly***Asn		529	PR	16-JUN-1999;	99US-0139452P.
					PR	16-JUN-1999;	99US-0139453P.
Db	1339	GTGAGTGTCTTCACTCTCGTGTTCGTCATATTTTGGGTGCCAGACAAAGGGAAG-AAC		1397	PR	17-JUN-1999;	99US-0139492P.
					PR	18-JUN-1999;	99US-0139454P.
Qy	530	SerArgGlyAspThrIlePheValSerLeu		539	PR	18-JUN-1999;	99US-0139455P.
					PR	18-JUN-1999;	99US-0139456P.
Db	1398	TCTGGAGGAGATACAGTTTTCGTTCTGCTA		1427	PR	18-JUN-1999;	99US-0139457P.
					PR	18-JUN-1999;	99US-0139458P.
					PR	18-JUN-1999;	99US-0139459P.
					PR	18-JUN-1999;	99US-0139460P.
					PR	18-JUN-1999;	99US-0139461P.
					PR	18-JUN-1999;	99US-0139462P.
					PR	18-JUN-1999;	99US-0139463P.
					PR	18-JUN-1999;	99US-0139750P.
					PR	18-JUN-1999;	99US-0139763P.
					PR	21-JUN-1999;	99US-0139817P.
					PR	22-JUN-1999;	99US-0139899P.
					PR	23-JUN-1999;	99US-0140353P.
					PR	23-JUN-1999;	99US-0140354P.
					PR	24-JUN-1999;	99US-0140695P.
					PR	28-JUN-1999;	99US-0140823P.
					PR	29-JUN-1999;	99US-0140991P.
					PR	30-JUN-1999;	99US-0141287P.
					PR	01-JUL-1999;	99US-0141842P.
					PR	01-JUL-1999;	99US-0142154P.
					PR	02-JUL-1999;	99US-0142055P.
					PR	06-JUL-1999;	99US-0142390P.
					PR	08-JUL-1999;	99US-0142803P.
					PR	09-JUL-1999;	99US-0142920P.
					PR	12-JUL-1999;	99US-0142977P.
					PR	13-JUL-1999;	99US-0143542P.
					PR	14-JUL-1999;	99US-0143624P.
					PR	15-JUL-1999;	99US-0144005P.
					PR	16-JUL-1999;	99US-0144085P.
					PR	16-JUL-1999;	99US-0144086P.
					PR	19-JUL-1999;	99US-0144325P.
					PR	19-JUL-1999;	99US-0144331P.
					PR	19-JUL-1999;	99US-0144332P.
					PR	19-JUL-1999;	99US-0144333P.
					PR	19-JUL-1999;	99US-0144334P.
					PR	19-JUL-1999;	99US-0144335P.
					PR	20-JUL-1999;	99US-0144352P.
					PR	20-JUL-1999;	99US-0144632P.
					PR	20-JUL-1999;	99US-0144884P.
					PR	21-JUL-1999;	99US-0144814P.
					PR	21-JUL-1999;	99US-0145086P.
					PR	21-JUL-1999;	99US-0145088P.
					PR	22-JUL-1999;	99US-0145085P.
					PR	22-JUL-1999;	99US-0145087P.
					PR	22-JUL-1999;	99US-0145089P.
					PR	22-JUL-1999;	99US-0145192P.
					PR	23-JUL-1999;	99US-0145145P.
					PR	23-JUL-1999;	99US-0145218P.
					PR	23-JUL-1999;	99US-0145224P.
					PR	26-JUL-1999;	99US-0145276P.
					PR	27-JUL-1999;	99US-0145913P.
					PR	27-JUL-1999;	99US-0145918P.
					PR	27-JUL-1999;	99US-0145919P.
					PR	28-JUL-1999;	99US-0145951P.
					PR	02-AUG-1999;	99US-0146386P.
					PR	02-AUG-1999;	99US-0146388P.
					PR	02-AUG-1999;	99US-0146389P.

RESULT 4

ID AAC36319 standard; DNA; 1741 BP.

XX AAC36319;

XX

DT 17-OCT-2000 (first entry)

XX

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 13356.

XX

KW Hybridisation assay; genetic mapping; gene expression control;

KW protein identification; signal transduction pathway; metabolic pathway;

KW promoter; termination sequence; ss.

XX

OS Arabidopsis thaliana.

XX

PN EP1033405-A2.

XX

PD 06-SEP-2000.

XX

PF 25-FEB-2000; 2000EP-00301439.

XX

PR 25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-0123180P.

PR 09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125788P.

PR 25-MAR-1999; 99US-0126264P.

PR 29-MAR-1999; 99US-0126785P.

PR 01-APR-1999; 99US-0127462P.

PR 06-APR-1999; 99US-0128234P.

PR 08-APR-1999; 99US-0128714P.

PR 16-APR-1999; 99US-0129845P.

PR 19-APR-1999; 99US-0130077P.

PR 21-APR-1999; 99US-0130449P.

PR 23-APR-1999; 99US-0130510P.

PR 23-APR-1999; 99US-0130891P.

PR 28-APR-1999; 99US-0131449P.

PR 30-APR-1999; 99US-0132048P.

PR 30-APR-1999; 99US-0132407P.

PR 04-MAY-1999; 99US-0132484P.

PR 05-MAY-1999; 99US-0132485P.

PR 06-MAY-1999; 99US-0132486P.

PR 06-MAY-1999; 99US-0132487P.

PR 07-MAY-1999; 99US-0132863P.

PR 11-MAY-1999; 99US-0134256P.

PR 14-MAY-1999; 99US-0134218P.

PR 14-MAY-1999; 99US-0134219P.

PR 14-MAY-1999; 99US-0134221P.

PR 14-MAY-1999; 99US-0134370P.

PR 14-MAY-1999; 99US-0134768P.

PR 18-MAY-1999; 99US-0134941P.

PR 19-MAY-1999; 99US-0135124P.

PR 20-MAY-1999; 99US-0135353P.

PR 21-MAY-1999; 99US-0135629P.

PR 24-MAY-1999;

PR	03-AUG-1999;	99US-0147038P.
PR	04-AUG-1999;	99US-0147204P.
PR	04-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147192P.
PR	05-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161992P.

PR	28-OCT-1999;	99US-0161993P.
PR	29-OCT-1999;	99US-0162142P.

Alignment Scores:
Pred. No.: 2,52e-76 Length: 1741
Score: 951.50 Matches: 181
Percent Similarity: 66.12% Conservative: 102
Best Local Similarity: 42.29% Mismatches: 140
Query Match: 34:24% Indels: 5
DB: 3 Gaps: 2

US-10-051-909-36 (1-553) x AAC36319 (1-1741)

QY	100	PheLeuCysThrLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGlyPhe	119
Db	287	TATCTCAGCACATTCGTTGCGGCTCTGTGGTCCTTTTGCCTTCGTTCTTGTGCGGATAC	346
QY	120	SerSerProThrGlnAspAlaMetValArgAspLeuAsnLeuSerIleSerGlyPheSer	139
Db	347	TCTTCACCTGCTCAGGCTGCAATTAGGAATGACCTTTTCATTGACTATAGCTGAGTTTCA	406
QY	140	AlaPheGlySerLeuSerAsnValGlyGlyMetValGlyAlaIleAlaSerGlyGlnMet	159
Db	407	CTCTTTGGTTCTTTACTAACTTTTGGCGCAATGATCGGTGCTATAACAAGCGGCCCTATA	466
QY	160	AlaGluTyrIleGlyArgLysGlySerLeuMetIleAlaAlaIleProAsnIleIleGly	179
Db	467	GCTGATTAGTTGGAAGAAAGGGCGGATGAGAGTTTCTCTGTCATTTTGTGTAGTGGG	526
QY	180	TrpLeuAlaIleSerPheAlaLysAspAlaSerPheLeuTyrMetGlyArgLeuLeuGlu	199
Db	527	TGGCTAGCAATCATCTTTGCCAAGGGAGTAGTGGCTCTGGACCTCGGAAGACTGGCAACG	586
QY	200	GlyPheGlyValGlyIleIleSerTyrThrValProValTyrIleAlaGluIleSerPro	219
Db	587	GGATATGGAATGGGAGCATTTTCTATGTGGTGCCCAATCTTTATAGCAGAAATTCACCT	646
QY	220	GlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIlePhe	239
Db	647	AAAACCTTCAGAGGGGCTCTAACCCACACTGAACACAGATTCTGTACTGGAGTGTCTG	706
QY	240	LeuAlaTyrLeuLeuGlyMetPheIleProTyrArgLeuLeuAlaValIleGlyAlaLeu	259
Db	707	GTTTCCTTCATCATAGGCACACTAGTAGCTGACGTGGAGAGTCTTGGCATTATAGGAATCATC	766
QY	260	ProCysThrMetLeuIleProGlyLeuPhePheIleProGluSerProArgTrpLeuAla	279
Db	767	CCATGCGCTGCCTCCTTCCCTTGGCCCTCTTTTATCCCTGAGTCTCCAAGATGGCTGGCA	826
QY	280	LysMetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPheGluThr	299
Db	827	AAAGTGGGGCGTGATACGGAGTTTGAAGCTGCACTGAGGAAGCTCCGTGGGAAGAAGGCT	886
QY	300	AspIleThrThrGluValAsnAspIleLysArgAlaValAlaSerSerSerLysArgThr	319
Db	887	GATATTTCGGAGGAGGCAGAGATCCAGGATTATATCGAAACTCTGGAAAGGCTACCG	946
QY	320	ThrIleSerPheGlnGluLeuAsnGlnLysLysTyrArgThrProLeuLeuLeuGlyIle	339
Db	947	AAAGCCAAGATGCTGGATTGTTTTCAGAGGAGATACATACGCTCTGTCTTATAGCTTTTC	1006
QY	340	GlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSer	359
Db	1007	GGGTTGATGGTGTTCAGCAGTTTGGAGGAATCAACGGAATATGTTTCTACACAAGCTCG	1066
QY	360	IlePheLysAlaAlaGlyValThrAsnSerAspIleuAlaThrCysSerLeuGlyAlaIle	379
Db	1067	ATATTGAGCAAGCAGGTTTCCC--ACAAGACTTGGGATGATATATATATGCTGTTCTT	1123
QY	380	GlnValLeuAlaThrGlyValThrThrTrpLeuLeuAspArgAlaGlyArgIleLeu	399
Db	1124	CAGGTGGTAATCACTGCGCTTAATGCACCCGATAGTTGACAGAGCCGGAAGAAACCATTG	1183

PR	23-JUL-1999;	99US-0145224P.
PR	26-JUL-1999;	99US-0145276P.
PR	27-JUL-1999;	99US-0145913P.
PR	27-JUL-1999;	99US-0145918P.
PR	27-JUL-1999;	99US-0145919P.
PR	28-JUL-1999;	99US-0145951P.
PR	02-AUG-1999;	99US-0146386P.
PR	02-AUG-1999;	99US-0146388P.
PR	02-AUG-1999;	99US-0146389P.
PR	03-AUG-1999;	99US-0147038P.
PR	04-AUG-1999;	99US-0147204P.
PR	04-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147192P.
PR	05-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.

PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161992P.
PR	28-OCT-1999;	99US-0161993P.
PR	29-OCT-1999;	99US-0162142P.

Alignment Scores:

Pred. No.:

1.28e-75

Length:

1683

Score:

943.50

Matches:

180

Percent Similarity:

65.89%

Conservative:

102

Best Local Similarity:

42.06%

Mismatches:

141

Query Match:

33.95%

Indels:

5

DB:

3

Gaps:

2

US-10-051-909-36 (1-553) x AAC39099 (1-1683)

Qy	100	PheLeuCysThrLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGlyPhe	119
Db	207	TATCTCAGCACATTCGTTGGGTCTGTTGGTTCGTTTCGGTTCGGTTCGTTTCGGGATAC	266
Qy	120	SerSerProThrGlnAspAlaMetValArgAspLeuAsnLeuSerIleSerGluPheSer	139
Db	267	TCCTCACCTGCTCAGGCTGCAATTAGGAATGACCTTTCATTGACTATAGCTGAGTTTCA	326
Qy	140	AlaPheGlySerLeuSerAsnValGlyGlyMetValGlyAlaIleAlaSerGlyGlnMet	159
Db	327	CTCTTTGGTTCCTTACTAACTTTTGGCGCAATGATCGGTGCTATACAAGCGGCCTATA	386
Qy	160	AlaGluTyrIleGlyArgGlySerLeuMetIleAlaAlaIleProAsnIleIleGly	179
Db	387	GCTGATTTAGTTGGAAGAAAGGGCGATGAGAGTTTCCTCTGCATTGTGTAGTCGGG	446
Qy	180	TrpLeuAlaIleSerPheAlaLysAspAlaSerPheLeuTyrMetGlyArgLeuLeuGlu	199
Db	447	TGGCTAGCAATCATCTTTGCCAAGGGAGTAGTGGCTCTGGACCTCGGAAGACTGGCAACG	506
Qy	200	GlyPheGlyValGlyIleIleSerTyrThrValProValTyrIleAlaGluIleSerPro	219
Db	507	GGATATGGAATGGGAGCATTTTCCTATGTCGTCCTAATCTTTATAGCAGAAATTCACCT	566
Qy	220	GlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIlePhe	239
Db	567	AAACTTTCAGAGGGGCTCTAACACACTGAACACAGATTCCTGATCTGCAGTGGAGTGTG	626
Qy	240	LeuAlaTyrLeuLeuGlyMetPheIleProTrpArgLeuLeuAlaValIleGlyAlaLeu	259
Db	627	GTTTCCTTCATCATTAGGCACACTAGTGACGTGGAGAGTCTTGGCATTATAGGAATCATC	686
Qy	260	ProCysThrMetLeuIleProGlyLeuPhePheIleProGluSerProArgTrpLeuAla	279
Db	687	CCATGCGCTGCCCTCCTTCCTTGGCCTCTTTTATCCCTGAGTCTCCAGATGGCTGGCA	746
Qy	280	LysMetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPheGluThr	299
Db	747	AAAGTGGGGCGGTATACGGAGTTTGAAGCTGCACCTGAGGAAGCTCCGTGGGAAGAGGCT	806
Qy	300	AspIleThrThrGluValAsnAspIleLysArgAlaValAlaSerSerSerLysArgThr	319
Db	807	GATATTTCCGAGGAGGCAGACAGATCCAGGATTATATCGAAACTCTCGAAAGGCTACCG	866
Qy	320	ThrIleSerPheGlnGluLeuAsnGlnLysTyrArgThrProLeuLeuLeuGlyIle	339
Db	867	AAAGCCAAGATGCTGATTTGTTTCAGAGGAGATACATACGCTCTGTCTTATAGCTTTC	926
Qy	340	GlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSer	359
Db	927	GGGTTGATGGTGTTCAGCAGTTCGAGGGAATCAACGGAATATGTTTCTACACAAGCTCG	986

Db	685	GCTGATATTTCGGAGGAGGCAGCAGAGATCCAGGATTATATCGAACTCTCGAAAGGCTA	744	PR	29-MAR-1999;	99US-0126785P.
Qy	319	ThrThrIleSerPheGlnGluLeuAsnGlnLysLysTyrArgThrProLeuLeuLeuGly	338	PR	01-APR-1999;	99US-0127462P.
Dd	745	CCGAAAGCCAAGATGCTGGATTGTTTCAGAGGAGATACATACGCTCTGTTCTTATAGCT	804	PR	06-APR-1999;	99US-0128234P.
Qy	339	IleGlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSer	358	PR	08-APR-1999;	99US-0128714P.
Dd	805	TTCGGGTGATGGTGTTCAGCAGTTTGGAGGAATCAACGAATATGTTTCTACACAAGC	864	PR	16-APR-1999;	99US-0129845P.
Qy	359	SerIlePheLysAlaAlaGlyValThrAsnSerAspLeuAlaThrCysSerLeuGlyAla	378	PR	19-APR-1999;	99US-0130077P.
Dd	865	TCGATATTGAGCAAGCAGGTTTCCCGGAGGATGATGATGATGATGATGATGATGATGAT	921	PR	21-APR-1999;	99US-0130449P.
Qy	379	IleGlnValLeuAlaThrGlyValThrThrTrpLeuLeuAspArgAlaGlyArgArgIle	398	PR	23-APR-1999;	99US-0130510P.
Dd	922	CTTCAGGTGGTAATCACTGCGCTTAATGCACCGATAGTTGACAGAGCCGGAACCA	981	PR	23-APR-1999;	99US-0130891P.
Qy	399	LeuLeuIleIleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSerValValPhe	418	PR	28-APR-1999;	99US-0131449P.
Dd	982	TTGCTACTGGTTCTGCAACAGGGTTAGTGATAGGCTGTGTGATAGCAGCGGTTCTTTC	1041	PR	30-APR-1999;	99US-0132048P.
Qy	419	PheLeuLysAspAsnIleSerGlnAspSerAsnSerTyrTyrIleLeuThrMetIleSer	438	PR	30-APR-1999;	99US-0132407P.
Dd	1042	TATCTCAAGGTTTAC-----GACATGGCGCAGCAAGCAGTCCCAGTCTCGGCT	1089	PR	04-MAY-1999;	99US-0132484P.
Qy	439	LeuValGlyIleValSerPheValIleThrPheSerPheGlyMetGlyAlaIleProTrp	458	PR	05-MAY-1999;	99US-0132485P.
Dd	1090	GTTGTTGGTATAATGGTGTACATAGGATCGTTTTTCAGCAGGAATGGAGCAATGCGGTGG	1149	PR	06-MAY-1999;	99US-0132486P.
Qy	459	LeuMetMetSerGluIleLeuProValSerIleLysSerLeuGlyGlySerIleAlaThr	478	PR	07-MAY-1999;	99US-0132863P.
Dd	1150	GTCGTCATGTCTGAGATATTTCCCATAAACATAAAGAGTAGCAGGAGGATGGCGACG	1209	PR	11-MAY-1999;	99US-0134256P.
Qy	479	LeuAlaAsnTrpLeuThrSerPheAlaIleThrMetThrThrAsnLeuMetLeuThrTrp	498	PR	14-MAY-1999;	99US-0134218P.
Dd	1210	CTGCTGAACCTGTTTGGAGCGTGGCTGTTTCTTACACTTCAACTTCCCTCATGTCCTGG	1269	PR	14-MAY-1999;	99US-0134219P.
Qy	499	SerValGlyGlyThrPheLeuSerTyrMetValValSerAlaPheThrIleValPheVal	518	PR	14-MAY-1999;	99US-0134221P.
Dd	1270	AGCTCTTACGGAACCTTTCCTCATTTACGCTGCCATCAACGCGTGGCCATTGTTCTCGTC	1329	PR	14-MAY-1999;	99US-0134370P.
Qy	519	ValLeuTrpValProGluThrLysGly	527	PR	18-MAY-1999;	99US-0134768P.
Dd	1330	ATTGCCATCGTGCCTGAGACAAAGGG	1356	PR	19-MAY-1999;	99US-0134941P.
RESULT 7						99US-0135124P.
AAC45298						99US-0135353P.
ID	AAC45298 standard; DNA; 1685 BP.					99US-0135629P.
XX						99US-0136021P.
AC	AAC45298;					99US-0136392P.
XX						99US-0136782P.
DT	18-OCT-2000 (first entry)					99US-0137222P.
XX						99US-0137528P.
DE	Arabidopsis thaliana DNA fragment SEQ ID NO: 46007.					99US-0137502P.
XX						99US-0137724P.
KW	Hybridisation assay; genetic mapping; gene expression control;					99US-0138094P.
KW	protein identification; signal transduction pathway; metabolic pathway;					99US-0138540P.
XX	promoter; termination sequence; ss.					99US-0138847P.
OS	Arabidopsis thaliana.					99US-0139119P.
PN	EP1033405-A2.					99US-0139452P.
XX						99US-0139453P.
PD	06-SEP-2000.					99US-0139457P.
XX						99US-0139458P.
PF	25-FEB-2000; 2000EP-00301439.					99US-0139459P.
XX						99US-0139460P.
PR	25-FEB-1999; 99US-0121825P.					99US-0139461P.
PR	05-MAR-1999; 99US-0123180P.					99US-0139462P.
PR	09-MAR-1999; 99US-0123548P.					99US-0139463P.
PR	23-MAR-1999; 99US-0125788P.					99US-0139750P.
PR	25-MAR-1999; 99US-0126264P.					99US-0139763P.

PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144362P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.

PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.
PR 29-OCT-1999; 99US-0162143P.

Alignment Scores:
Pred. No.: 9.39e-74
Score: 923.00
Percent Similarity: 65.58%
Best Local Similarity: 42.09%
Query Match: 33.21%
DB: 3

US-10-051-909-36 (1-553) x AAC45298 (1-1685)

Qy 100 PheLeuCysThrLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGlyPhe 119
Db 207 TATCTCAGCACATTCGTTGCGGCTCTGTTGCGTTCTTTTTCGTTCTTGTGCGGGATAC 266

Qy 120 SerSerProThrGlnAspAlaMetValArgAspLeuAsnLeuSerIleSerGluPheSer 139
Db 267 TCTTCACCTGCTCAGGCTGCAATTAGGAATGACCTTTTCATTGACTATAGCTGAGTTTCA 326

Qy 140 AlaPheGlySerLeuSerAsnValGlyGlyMetValGlyAlaIleAlaSerGlyGlnMet 159
Db 327 CTCCTTTGGTTCTTTACTAACTTTTGGCGCAATGATCGGTGTCTATAACAAGCGGCTATA 386

Qy 160 AlaGluTyrIleGlyArgLysGlySerLeuMetIleAlaAlaIleProAsnIleIleGly 179
Db 387 GCTGATTAGTTGGAAGAAAGGGGCGATGAGAGTTTCTCTGCAATTTGTGTAGTCGGG 446

Qy 180 TrpLeuAlaIleSerPheAlaLysAspAlaSerPheLeuTyrMetGlyArgLeuGlu 199
Db 447 TGGCTAGCAATCATCTTTGCCAAGGGAGTAGTGGCTCTGGACCTTGGAAAGACTGGCAACG 506

Qy 200 GlyPheGlyValGlyIleIleSerTyrThrValProValTyrIleAlaGluIleSerPro 219
Db 507 GGATATGGAATGGGAGCATTTTCTCTATGTGGTGCCAAATCTTTATAGCAGAAATTCACCT 566

Qy 220 GlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeu---Ser-ValThrPheGlyI 238
Db 567 AAAACCTTCAGAGGGGCTCTAACCACTACACTGAACACAGACAGATTTCTGATCTGCAC 626

Qy 238 ePheLeuAlaTyrLeuLeuGlyMetPheIleProTyrArgLeuLeuAlaValIleGlyAl 258
Db 627 GTCGGTTTCTTCATCATAGGCACACTAGTACGTCGAGAGTCTTGGCATTATAGGAAT 686

Qy 258 aLeuProCysThrMetLeuIleProGlyLeuPhePheIleProGluSerProArgTrpLe 278

Db	687	CATCCCATGGCTGCCTCCTTCCCTTGGCCTCTTTTATCCCTGAGTCTCCAAGATGGCT	746	PD	06-SEP-2000.
QY	278	uAlaLysMetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPheG1	298	XX	
Db	747	GGCAAAAGTGGGGCGTCATACGGAGTTTGAAGTGCACCTGAGGAAGCTCCGTGGGAAGAA	806	PF	25-FEB-2000; 2000EP-00301439.
QY	298	uThrAspIleThrThrGluValAsnAspIleLysArgAlaValAlaSerSerLysAr	318	XX	25-FEB-1999; 99US-0121825P.
Db	807	GGCTGATATTTCGGAGGAGGCAGACAGATCCAGGATTATATCGAAACTCTGGAAGGCT	866	PR	05-MAR-1999; 99US-0123180P.
QY	318	gThrThrIleSerPheGlnGluLeuAsnGlnLysLysTyrArgThrProLeuLeuG1	338	PR	09-MAR-1999; 99US-0123548P.
Db	867	ACCGAAAGCCAAAGATGCTGGATTGTTTCAGAGGAGATACATACGCTCTGTTCTTATAGC	926	PR	23-MAR-1999; 99US-0125788P.
QY	338	yIleGlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSe	358	PR	25-MAR-1999; 99US-0126264P.
Db	927	TTTCGGGTGTGATGGTGTTCAGCAGATTTCGGAGGAATCAACGGAATATGTTTCTACACAAG	986	PR	99US-0126785P.
QY	358	rSerIlePheLysAlaAlaGlyValThrAsnSerAspLeuAlaThrCysSerLeuGlyAl	378	PR	01-APR-1999; 99US-0127462P.
Db	987	CTCGATATTTGAGCAAGCAGGTTTTCCTCC---ACAAGACTTGGGATGATAATATATGCTGT	1043	PR	06-APR-1999; 99US-0128234P.
QY	378	aIleGlnValLeuAlaThrGlyValThrThrTrpLeuLeuAspArgAlaGlyArgArgI1	398	PR	08-APR-1999; 99US-0128714P.
Db	1044	TCTTCAGGTGGTAATCACTGCGCTTAATGCACCGATGTTGACAGAGCCGGAAGAAACC	1103	PR	16-APR-1999; 99US-0129845P.
QY	398	eLeuLeuIleIleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSerValValPh	418	PR	19-APR-1999; 99US-0130077P.
Db	1104	ATTGCTACTGTTTCTGCAACAGCGGTTAGTGATAGCTGTTTGTATAGCAGCGTTTCTTT	1163	PR	21-APR-1999; 99US-0130449P.
QY	418	ePheLeuLysAspAsnIleSerGlnAspSerAsnSerTyrTyrIleLeuThrMetIleSe	438	PR	23-APR-1999; 99US-0130510P.
Db	1164	CTATCTCAAGGTTTCAC-----GACATGGCGCACGAAGCAGTCCAGTCCCTGGC	1211	PR	23-APR-1999; 99US-0130891P.
QY	438	rLeuValGlyIleValSerPheValIleThrPheSerPheGlyMetGlyAlaIleProTr	458	PR	28-APR-1999; 99US-0131449P.
Db	1212	TGTTGTTGGTATAATGGTGATAGGATCGTTTTCAGCAGGAATGGGAGCAATGCCGTG	1271	PR	30-APR-1999; 99US-0132048P.
QY	458	pLeuMetMetSerGluIleLeuProValSerIleLysSerLeuGlyGlySerIleAlaTh	478	PR	30-APR-1999; 99US-0132407P.
Db	1272	GGTGGTCATGCTGAGATATTCCCATAAACATAAAGGAGTAGCAGGAGGCATGCCGAC	1331	PR	04-MAY-1999; 99US-0132484P.
QY	478	rLeuAlaAsnTrpLeuThrSerPheAlaIleThrMetThrThrAsnLeuMetLeuThrTr	498	PR	05-MAY-1999; 99US-0132485P.
Db	1332	GCTGGTGAACCTGGTTTGGACCGTGGGCTGTTCTTCATCACTTTCAACTTCCTCATGTCCTG	1391	PR	06-MAY-1999; 99US-0132486P.
QY	498	pSerValGlyGlyThrPheLeuSerTyrMetValValSerAlaPheThrIleValPheVa	518	PR	07-MAY-1999; 99US-0132863P.
Db	1392	GAGCTCTTACGGAACCTTTCCTCATTTACGCTGCCATCAACGGCGCTGGCCATGTCTTCGT	1451	PR	11-MAY-1999; 99US-0132866P.
QY	518	lValLeuTrpValProGluThrLysGly	527	PR	14-MAY-1999; 99US-0134218P.
Db	1452	CATTGCCATCGTGCCTGAGACAAAAGGG	1479	PR	14-MAY-1999; 99US-0134219P.
RESULT 8					PR 14-MAY-1999; 99US-0134370P.
AAC45290					PR 18-MAY-1999; 99US-0134768P.
ID AAC45290 standard; DNA; 1743 BP.					PR 18-MAY-1999; 99US-0134941P.
XX					PR 20-MAY-1999; 99US-0135124P.
AC AAC45290;					PR 21-MAY-1999; 99US-0135353P.
XX					PR 24-MAY-1999; 99US-0135629P.
DT 18-OCT-2000 (first entry)					PR 25-MAY-1999; 99US-0136021P.
XX					PR 27-MAY-1999; 99US-0136392P.
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 45977.					PR 28-MAY-1999; 99US-0136782P.
XX					PR 01-JUN-1999; 99US-0137222P.
KW Hybridisation assay; genetic mapping; gene expression control;					PR 03-JUN-1999; 99US-0137528P.
XW protein identification; signal transduction pathway; metabolic pathway;					PR 04-JUN-1999; 99US-0137502P.
KW promoter; termination sequence; ss.					PR 07-JUN-1999; 99US-0137724P.
XX					PR 08-JUN-1999; 99US-0138094P.
OS Arabidopsis thaliana.					PR 10-JUN-1999; 99US-0138540P.
XX					PR 10-JUN-1999; 99US-0138847P.
PN EP1033405-A2.					PR 14-JUN-1999; 99US-0139119P.
XX					PR 16-JUN-1999; 99US-0139452P.

Hybridisation assay; genetic mapping; gene expression control;
protein identification; signal transduction pathway; metabolic pathway;
promoter; termination sequence; ss.

PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 05-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 06-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.

PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:

Pred. No.: 9.87e-74 Length: 1743
Score: 923.00 Matches: 181
Percent Similarity: 65.58% Conservative: 101
Best Local Similarity: 42.09% Mismatches: 141
Query Match: 33.21% Indels: 7
DB: 3 Gaps: 3

US-10-051-909-36 (1-553) x AAC45290 (1-1743)

QY 100 PheLeuCysThrLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGlyPhe 119
Db 287 TATCTCAGCACATTCGTTGCGGTCCTGCGTTCTTTTTCGTTTCGTTCTTGTGCGGGATAC 346
QY 120 SerSerProThrGlnAspAlaMetValArgAspLeuAsnLeuSerIleSerGlyPheSer 139
Db 347 TCTTCACCTGCTCAGGCTGCAATTAGGAATGACCTTTTCATTGACTATAGCTGAGTTTTC 406
QY 140 AlapheGlySerLeuSerAsnValGlyGlyMetValGlyAlaIleAlaSerGlyGlnMet 159
Db 407 CTCTTTGGTTCTTTACTAACTTTTGGCGCAATGATCGGTGCTATAACAAGCGGGCTATA 466
QY 160 AlaGluTyriIleGlyArgLysGlySerLeuMetIleAlaAlaIleProAsnIleIleGly 179
Db 467 GCTGATTAGTTGGAAGAAAGGGGCGCATGAGAGTTTCTCTGCAATTTGTGTAGTCGGG 526
QY 180 TrpLeuAlaIleSerPheAlaLysAspAlaSerPheLeuTyMetGlyArgLeuLeuGlu 199
Db 527 TGGCTAGCAATCATCTTTGCCAAGGGAGTAGTGGCTCTGGACCTTGAAGACTGGCAACG 586
QY 200 GlyPheGlyValGlyIleIleSerTyThrValProValTyIleAlaGluIleSerPro 219
Db 587 GGATATGGAATGGAGCATTTTTCCTATGTGGTGCCCAATCTTATATAGCAGAAATTGCACCT 646

QY 220 GlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeu---Ser-ValThrPheGlyI1 238
Db 647 AAAAATTTTCAGAGGGGCTTAACACACTGAACACAGACAGATTCTGATCTGCACCTGGAGT 706
QY 238 ePheLeuAlaTyLeuLeuGlyMetPheIleProTrpArgLeuLeuAlaValIleGlyAl 258
Db 707 GTCGGTTTCCTCATCATAGGCACACTAGTAGCGTGGAGAGTCTTGGCATTATAGGAAT 766
QY 258 aLeuProCysThrMetLeuIleProGlyLeuPhePheIleProGlySerProArgTrpLe 278
Db 767 CATCCCATGCGCTGCCTCTTCTTCTTGGCCTCTTTTATCCCTGAGTCTCCAAGATGGCT 826
QY 278 uAlaLysMetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPheG1 298
Db 827 GGCAAAAGTGGGGCGGTATACGGAGTTTGAAGCTGCACTGAGGAAGCTCCGTGGGAAGAA 886
QY 298 uThrAspIleThrThrGluValAsnAspIleLysArgAlaValAlaSerSerLysAr 318
Db 887 GGCTGATATTTCGGAGGAGGCAGCAGAGATCCAGGATTATATCGAAACTCTGGAAGGCT 946
QY 318 gThrThrIleSerPheGlnGluLeuAsnGlnLysLysTyArgThrProLeuLeuGlu 338
Db 947 ACCGAAAGCCCAAGATGTGGATTGTTTCAGAGAGATACATACGCTCTGTCTTATAGC 1006
QY 338 yIleGlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyAlaSe 358
Db 1007 TTTCGGGTGATGGTGTTCAGCAGATTTCAGGAGGATCAACGGAATATGTCTACACAAG 1066
QY 358 rSerIlePheLysAlaAlaGlyValThrAsnSerAspLeuAlaThrCysSerLeuGlyAl 378
Db 1067 CTCGATATTGAGCAAGCAGGTTTTCCTCC---ACAGACTTGGGATGATATATATGCTGT 1123
QY 378 aIleGlnValLeuAlaThrGlyValThrThrTrpLeuLeuAspArgAlaGlyArgI1 398
Db 1124 TCTTCAGGTGGTAATCACTCGGCTTAATGCACCGATAGTTGACAGAGCCGGAAGAAACC 1183
QY 398 eLeuLeuIleIleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSerValPh 418
Db 1184 ATTGCTACTGTTTCTGCAACAGGGTTAGTGATAGCTGTTTGATAGCAGCGGTTCTTT 1243
QY 418 ePheLeuLysAspAsnIleSerGlnAspSerAsnSerTyTyIleLeuThrMetIleSe 438
Db 1244 CTATCTCAAGGTTTCAC-----GACATGGCGCACGAAGCAGTCCCGCTCTGGC 1291
QY 438 rLeuValGlyIleValSerPheValIleThrPheSerPheGlyMetGlyAlaIleProTr 458
Db 1292 TGTTGTTGGTATATATGTTGTACATAGGATCGTTTTTCAGCAGGAATGGGAGCAATGCCGTG 1351
QY 458 pLeuMetMetSerGluIleLeuProValSerIleLysSerLeuGlyGlySerIleAlaTh 478
Db 1352 GGTTGTCATGTCGTGAGATATTTCACATAAACAATAAAGAGTAGCAGGAGGATGGCGAC 1411
QY 478 rLeuAlaAsnTrpLeuThrSerPheAlaIleThrMetThrThrAsnLeuMetLeuThrTr 498
Db 1412 GCTGGTGAACGGTTTGGAGCGGTGGCTGTTTCTTACACTTTCACACTTCTCTCATGCTCTG 1471
QY 498 pSerValGlyGlyThrPheLeuSerTyMetValValSerAlaPheThrIleValPheVa 518
Db 1472 GAGCTCTTACGGAACCTTTCCTCATTTACGCTGCCATCACCGGCTGGCCATTGTCTTCGT 1531
QY 518 lValleuTrpValProGluThrLysGly 527
Db 1532 CATTCCTATCGTGCCTGAGACAAAAGGG 1559

RESULT 9

ABZ14526

ID ABZ14526 standard; DNA; 1473 BP.

XX AC ABZ14526;

XX AC ABZ14526;

DT 21-JAN-2003 (first entry)

XX Arabidopsis thaliana stress regulated gene SEQ ID NO 2331.

XX Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
KW Arabidopsis thaliana.
OS Arabidopsis thaliana.
XX WO200216655-A2.
XX 28-FEB-2002.
XX 24-AUG-2001; 2001WO-US026685.
XX 24-AUG-2000; 2000US-0227866P.
XX 26-JAN-2001; 2001US-0264647P.
XX 22-JUN-2001; 2001US-0300111P.
XX (SCRI) SCRIPPS RES INST.
XX (SYGN) SYNGENTA PARTICIPATIONS AG.
XX Harper JF, Kreps J, Wang X, Zhu T;
XX WPI; 2002-304127/34.
XX Identifying a stress condition to which a plant cell has been exposed and
XX producing plants with increased tolerance to these abiotic stresses.
XX Claim 144; SEQ ID NO 2331; 577pp + Sequence Listing; English.
XX The invention relates to identifying a stress condition to which a plant
XX cell has been exposed, comprising: (a) contacting nucleic acid
XX representative of expressed polynucleotides in the plant cell with an
XX array or probes representative of the plant cell genome; and (b)
XX detecting a profile of expressed polynucleotides in the plant cell
XX characteristic of a stress response. The method is useful in the
XX production of transgenic plants, cells and seeds and in producing plants
XX with increased tolerance to abiotic stress. The present sequence is that
XX of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
XX in methods of the invention. Note: The sequence data for this patent is
XX not represented in the printed specification but is based on sequence
XX information supplied to Derwent by the European Patent Office
XX
SQ Sequence 1473 BP; 372 A; 292 C; 352 G; 457 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.3e-73 Length: 1473
Score: 920.50 Matches: 186
Percent Similarity: 61.69% Conservative: 91
Best Local Similarity: 41.43% Mismatches: 159
Query Match: 33.12% Indels: 13
DB: 6 Gaps: 3

US-10-051-909-36 (1-553) x ABZ14526 (1-1473)

QY 98 SerAlaPheLeuCysThrLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSer 117
Db 151 TCTGTTTCTCAGCACCTTTGTGCGCGTATCAGGCTCCTTCTGTACCGGTTGTGGCGTT 210
QY 118 GlyPheSerSerProThrGlnAppAlaMetValArgAspLeuAsnLeuSerIleSerGlu 137
Db 211 GGTTTTTCATCGGGTGCACAGCAGGGATTACCAAGATTATCTCTCTCCGTTGCAGAA 270
QY 138 PheSerAlaPheGlySerLeuSerAsnValGlyGlyMetValGlyAlaIleAlaSerGly 157
Db 271 TACTCAATGTTCCGGTTCGATCTTTCACATTAGGAGGCTTGTATCGGTGCAGTATTTCAGCGGT 330
QY 158 GlnMetAlaGluTyIleGlyArgLysGlySerLeuMetIleAlaAlaIleProAsnIle 177
Db 331 AAAGTCGCTGATGTTCTTGGGAAGAAACGACGATGTTGTTTTCGGAATTCTTCTGTATC 390
QY 178 IleGlyTrpLeuAlaIleSerPheAlaLysAspAlaSerPheLeuTyMetGlyArgLeu 197
Db 391 ACAGGCTGGCTTGTGTAGCATTTGGCTCAGAAATGCAATGTGGCTGGACTGTGGAAGATTG 450
QY 198 LeuGluGlyPheGlyValGlyIleIleSerTyThrValProValTyIleAlaGluIle 217

Db 451 TTACTTGAATCGGGTGGTATATTTAGCTACGTGATTCGGGTGATATATAGCCGAAATT 510
Qy 218 SerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGly 237
Db 511 GCACCTAAACATGTCGAGGATCGTTTGTGTCGCCAATCAGTTGATGCAAAATTCGCGA 570
Qy 238 IlePheLeuAlaTyrLeuLeuGlyMetPheIleProTrpArgLeuLeuAlaValIleGly 257
Db 571 ATTTCACTCTTCTTCATCATTTGGCAATTTTATTCATGGAGACTACTAACAGTAGTCGGA 630
Qy 258 AlaLeuProCysThrMetLeuIleProGlyLeuPhePheIleProGluSerProArgTrp 277
Db 631 TTGGTGCCATGTGTGTTCCACGTCTTTTGTATTATTTTCATCCCGAATCTCCAAGATGG 690
Qy 278 LeuAlaLysMetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPhe 297
Db 691 CTGGCGAAGTTAGTTCGTGTATAAAGATGCGCATCTTCGTTCACAGCCTTAGGGGATCT 750
Qy 298 GluThrAspIleThrGluValAsnAspIleLysArgAlaValAlaSerSerSerLys 317
Db 751 GACGTCGATATTTCTCGTGAAGCAACAACATAGAGATACCATTCATGACATGACAGAAAAC 810
Qy 318 ArgThrThrIleSerPheGlnGluLeuAsnGlnLysLysTyrArgThrProLeuLeuLeu 337
Db 811 GGTGGTGAACCTAAGATGTCGTAATTTGTTTCAGAGACGATACGGATATCCGTTAATTATC 870
Qy 338 GlyIleGlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAla 357
Db 871 GGAGTTGGTTTAAATGTTTTTGCAACAATGTGTGGGAGCTCCGGTGTACCTATTATGCT 930
Qy 358 SerSerIlePheLysAlaAlaGlyValThrAsnSerAspLeuAlaThrCysSerLeuGly 377
Db 931 AGTAGCCTCTTCAACAAAGGAGGATTTCCA--AGTGCTATTGGCACATCCGTAATAGCC 987
Qy 378 AlaIleGlnValLeuAlaThrGlyValThrThrTrpLeuLeuAspArgAlaGlyArgArg 397
Db 988 ACAATTATGTTCCAAAAGCAATGCTGGCAACAGTCCTAGTCGATAAAATGGGGAGAGA 1047
Qy 398 IleLeuLeuIleIleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSerValVal 417
Db 1048 ACGCTCCTAATGGCTTCTTGTCTCGCAATGGGTTTGAGTGTCTTCTCTTAAGTGT-- 1104
Qy 418 PhePheLeuLysAspAsnIleSerGlnAspSerAsnSerTyrTyrIleLeu----- 434
Db 1105 -----TCTTACGGTTTCCAGTCGTTTGGCATTCTTCCAGAACTC 1143
Qy 435 ---ThrMetIleSerLeuValGlyIleValSerPheValIleThrPheSerPheGlyMet 453
Db 1144 ACTCCCATCTTCACTTGCATCGGCGTCTTGGTGCACATTGTGTCATTGGCCATGGGAATG 1203
Qy 454 GlyAlaIleProTrpLeuMetMetSerGluIleLeuProValSerIleLysSerLeuGly 473
Db 1204 GGAGGACTACCATGGATTATTAATGGCTGAGATATTTCCGATGAATGTGAAAGTGTACGCT 1263
Qy 474 GlySerIleAlaThrLeuAlaAsnTrpLeuThrSerPheAlaIleThrMetThrThrAsn 493
Db 1264 GGGACCTTAGTACTGTAAACCAATTGGTTATTTGGTTGGATTATCACATACACTTTCAAT 1323
Qy 494 LeuMetLeuThrTrpSerValGlyThrPheLeuSerTyrMetValValSerAlaPhe 513
Db 1324 TTTATGCTAGAAATGGAATGCATCAGGAATGTTCTCTCATCTCTCAATGGTCTCCGCGCAGT 1383
Qy 514 ThrIleValPheValValLeuTrpValProGluThrLysGly***AsnSerArgGlyAsp 533
Db 1384 TCGATCGTATTATATACCTTTTGGTACCTGAGACAAAGCCGATCACTTGAAGAAATA 1443
Qy 534 ThrIlePheValSerLeuSerIleGln 542
Db 1444 CAAGCACTGCTCAACAACCTCTGTGCAA 1470

ID AAC40459 standard; DNA; 1755 BP.
AC AAC40459;
XX 17-OCT-2000 (first entry)
DT Arabidopsis thaliana DNA fragment SEQ ID NO: 28375.
XX Arabidopsis thaliana.
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX Arabidopsis thaliana.
OS EP1033405-A2.
XX 06-SEP-2000.
PN 25-FEB-2000; 2000EP-00301439.
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 03-JUN-1999; 99US-0137222P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.

PR	23-AUG-1999;	99US-0149930P;
PR	25-AUG-1999;	99US-0150566P;
PR	26-AUG-1999;	99US-0150884P;
PR	27-AUG-1999;	99US-0151065P;
PR	27-AUG-1999;	99US-0151066P;
PR	27-AUG-1999;	99US-0151080P;
PR	30-AUG-1999;	99US-0151303P;
PR	31-AUG-1999;	99US-0151438P;
PR	01-SEP-1999;	99US-0151930P;
PR	07-SEP-1999;	99US-0152363P;
PR	10-SEP-1999;	99US-0153070P;
PR	13-SEP-1999;	99US-0153758P;
PR	15-SEP-1999;	99US-0154018P;
PR	16-SEP-1999;	99US-0154039P;
PR	20-SEP-1999;	99US-0154779P;
PR	22-SEP-1999;	99US-0155139P;
PR	23-SEP-1999;	99US-0155486P;
PR	24-SEP-1999;	99US-0155659P;
PR	28-SEP-1999;	99US-0156458P;
PR	29-SEP-1999;	99US-0156596P;
PR	04-OCT-1999;	99US-0157117P;
PR	05-OCT-1999;	99US-0157753P;
PR	06-OCT-1999;	99US-0157865P;
PR	07-OCT-1999;	99US-0158029P;
PR	08-OCT-1999;	99US-0158232P;
PR	12-OCT-1999;	99US-0158369P;
PR	13-OCT-1999;	99US-0159293P;
PR	13-OCT-1999;	99US-0159294P;
PR	13-OCT-1999;	99US-0159295P;
PR	14-OCT-1999;	99US-0159329P;
PR	14-OCT-1999;	99US-0159330P;
PR	14-OCT-1999;	99US-0159331P;
PR	14-OCT-1999;	99US-0159637P;
PR	14-OCT-1999;	99US-0159638P;
PR	18-OCT-1999;	99US-0159584P;
PR	21-OCT-1999;	99US-0160741P;
PR	21-OCT-1999;	99US-0160767P;
PR	21-OCT-1999;	99US-0160768P;
PR	21-OCT-1999;	99US-0160770P;
PR	21-OCT-1999;	99US-0160814P;
PR	21-OCT-1999;	99US-0160815P;
PR	22-OCT-1999;	99US-0160980P;
PR	22-OCT-1999;	99US-0160981P;
PR	22-OCT-1999;	99US-0160989P;
PR	25-OCT-1999;	99US-0161404P;
PR	25-OCT-1999;	99US-0161405P;
PR	25-OCT-1999;	99US-0161406P;
PR	26-OCT-1999;	99US-0161359P;
PR	26-OCT-1999;	99US-0161360P;
PR	26-OCT-1999;	99US-0161361P;
PR	28-OCT-1999;	99US-0161920P;
PR	28-OCT-1999;	99US-0161992P;
PR	28-OCT-1999;	99US-0161993P;
PR	29-OCT-1999;	99US-0162142P;

Alignment Scores
Pred. No.:
Score:
Percent Similarity
Best Local Similarity
Query Match:
DB:

URS-10-051-909-36 (1-553) x AAC40459 (1-1755)

Qy	94	GluserHisValSerAlaPheLeuCysThrLeuIleValAlaLeuGlyProIleGlnPhe	113
	188	GAATCATATTGATGGTTCCTCAGTACTTTTGGTCTGCTGGCTCCTTTGAGTTT	247
Db	114	GlyPheThrSerGlyPheSerSerProThrGlnAspAlaMetValArgAspLeuAsnLeu	133
	248	GGCTCTTGTTGGTACTCAGCGCTACTCAGTCACTATTAAGACAAGATCTCAATCTC	307

QY 134 SerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyGlyMetValGlyAla 153
Db 308 TCCTTGACAGAGTTCTCCATGTTTGGATCCATCTTAACCTATCGGTGCAATGCTGGTGCT 367
QY 154 IleAlaSerGlyGlnMetAlaGluTyrIleGlyArgLysGlySerLeuMetIleAlaAla 173
Db 368 GTATGAGTGGGAAATTTTCAGATTCTCCGGCCGAAAGGGCAATGAGGACGTCAGCT 427
QY 174 IleProAsnIleIleGlyTrpLeuAlaIleSerPheAlaLysAspAlaSerPheLeuTyr 193
Db 428 TGCTTCTGCATTACAGTTGGCTCGCTGCTTCTTCCACCAAGGGGCATTGTTACTTGAT 487
QY 194 MetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTyrTrpValProValTyr 213
Db 488 GTAGGAAGGTTCTTTACAGGATATGGAATTGGAGTTTCTTATGTTGCTCCTGTGTAC 547
QY 214 IleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSer 233
Db 548 ATTGCTGAGATATCTCCCAAGAATCTCCGAGGTGACTCACAACACTGAACCAACTCATG 607
QY 234 ValThrPheGlyIlePheLeuAlaTyrLeuLeuGlyMetPheIleProTrpArgLeuLeu 253
Db 608 ATTGTGATCGGCTCATCGGTTCTTCTTGTGCGGATCTCTCATTTCTTGGAACACTCTT 667
QY 254 AlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPhePheIleProGlu 273
Db 668 GCCCTAACCGTACTTGTCTCCCTGCTGCTTTGCTCTTTGGCTTGTGCTTATACCCGAA 727
QY 274 SerProArgTrpLeuAlaLysMetAsnLeuThrGluAspCysGluThrSerLeuGlnVal 293
Db 728 TCTCCTCGATGGCTGGCAAAAGCAGGCCCATGAGAAAGAGTTTCGGCTAGCCCTGCAAAAG 787
QY 294 LeuArgGlyPheGluThrAspIleThrThrGluValAsnAspIleLysArgAlaValAla 313
Db 788 CTGCGAGGAAAGATGCAGATATCAAAATGAAGCAGACGGTATTCAGTCTCGATTCAA 847
QY 314 SerSerSerLysArgThrThrIleSerPheGlnGluLeuAsnGlnLysLysTyrArgThr 333
Db 848 GCTCTAGAGATTCTTCCAAAGCAAGCAAGATCCCAAGACCTTGTGTCCAAGAAATATGGTCA 907
QY 334 ProLeuLeuGlyIleGlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyVal 353
Db 908 TCTGTATCATGTTGTTTCCCTGATGTTATCCCAACAGATTTTGTGAATCAATGGGATC 967
QY 354 LeuPheTyrAlaSerSerIlePheLysAlaAlaGlyValThrAsnSerAspLeuAlaThr 373
Db 968 GGATTCTACGCAAGTGAACGTTTGTAAAGCCGATTTACCTCTGGGAAACTAGGAACA 1027
QY 374 CysSerLeuGlyAlaIleGlnValLeuAlaThrGlyValThrThrTrpLeuLeuAspArg 393
Db 1028 A--TCGCTATCGTTGTTTCCAGGTGCCAATAAAGTCTTCTTGAACAATCTTGTATAGATAAA 1085
QY 394 AlaGlyArgArgIleLeuLeuIleIleSerThrSerGlyMetThrLeuCysLeuLeuAla 413
Db 1086 TCTGGACGAAGCCCAATATATGATTTAGCTGGTGGTATCTTCTTGGGATGCATATC 1145
QY 414 ValSerValValPhePheLeuLysAspAsnIleSerGlnAspSerAsnSerTyrTyrIle 433
Db 1146 ACAGGCACATCTTCTTACTCAAG-----GGACAGAGCTTGTGCTT 1187
QY 434 LeuThrMetIleSerLeuVal-----GlyIleValSerPheValIleThrPheSerPhe 451
Db 1188 GAATGGTCCCTTCTTAGCCGTTGGAGGTGATCTATCTATGATGCTGCTTCTTCTCATC 1247
QY 452 GlyMetGlyAlaIleProTrpLeuMetMetSerGluIleLeuProValSerIleLysSer 471
Db 1248 GGAATGGGACCTGTTCTTGGTGATAATGTCTGAGATATTTCCGATAAACGTAAGGGA 1307
QY 472 LeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeuThrSerPheAlaIleThrMetThr 491
Db 1308 ATCGCAGGAAGCTTAGTGGTACTAGTGAATTGGTCTGGTCTGGCTGTTCTTCTACT 1367
QY 492 ThrAsnLeuMetLeuThrTrpSerValGlyGlyThrPheLeuSerTyrMetValSer 511

Db 1368 TTCAACTTCTCATGAGCTGGAGCTCTCCAGGTACATCTATTGTACTCGCTTTTGCA 1427
QY 512 AlaPheThrIleValPheValValLeuTrpValProGluThrLysGly---***AsnSerA 531
Db 1428 GCTCGGACGATAATATTTGTGCGGAAGATGGTCCAGAGACGAAAGGGAAGACACTGGAA 1487
QY 531 rgGlyAspThrIlePheValSerLeuSerIleGlnArgGlnLeuGln---TrpLeuProG 550
Db 1488 GAGATCCAAGCTTGTATTTCGAAGAGAAACATAGAAAGAAATCAATCAATGTTATGGG 1547
QY 550 luCys 551
Db 1548 AATGT 1552
RESULT 11
ABZ13822
ID ABZ13822 standard; DNA; 1359 BP.
XX AC ABZ13822;
XX 21-JAN-2003 (first entry)
DE Arabidopsis thaliana stress regulated gene SEQ ID NO 1627.
KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX OS Arabidopsis thaliana.
XX PN WO200216655-A2.
XX 28-FEB-2002.
XX 24-AUG-2001; 2001WO-US026685.
PR 24-AUG-2000; 2000US-0227866P.
PR 26-JAN-2001; 2001US-0264647P.
PR 22-JUN-2001; 2001US-0300111P.
XX (SCRI) SCRIPPS RES INST.
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX Harper JF, Kreps J, Wang X, Zhu T;
XX WPI; 2002-304127/34.
PT Identifying a stress condition to which a plant cell has been exposed and
PT producing plants with increased tolerance to these abiotic stresses.
PS Claim 144; SEQ ID NO 1627; 577pp + Sequence Listing; English.
XX The invention relates to identifying a stress condition to which a plant
CC cell has been exposed, comprising: (a) contacting nucleic acid
CC representative of expressed polynucleotides in the plant cell with an
CC array or probes representative of the plant cell genome; and (b)
CC detecting a profile of expressed polynucleotides in the plant cell
CC characteristic of a stress response. The method is useful in the
CC production of transgenic plants, cells and seeds and in producing plants
CC with increased tolerance to abiotic stress. The present sequence is that
CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
CC in methods of the invention. Note: The sequence data for this patent is
CC not represented in the printed specification but is based on sequence
CC information supplied to Derwent by the European Patent Office
XX
SQ Sequence 1359 BP; 340 A; 238 C; 327 G; 454 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1.04e-71 Length: 1359
Score: 899.00 Matches: 188
Percent Similarity: 60.17% Conservative: 99
Best Local Similarity: 39.41% Mismatches: 149
Query Match: 32.35% Indels: 42
Gaps: 5
DB:

US-10-051-909-36 (1-553) x ABZ13822 (1-1359)		
QY	79	GlnSerSerPheAlaProGlyThrSerSerMetAlaValLeuArgGluSerHisValSer 98
Db	19	GAAGCTAATCTTGCTCCTGAAACTTCA-----CTTATTAAACAAGAGAACTCAAGACTCA 72
QY	99	Ala-----PheLeuCysThrLeuIleValAlaLeuGlyProIle 111
Db	73	TCGGCGACAATAACAACACTCTCTTGCTCACCACTTTTGTAGCCGTGTCGTTCCTTT 132
QY	112	GlnPheGlyPheThrSerGlyPheSerSerProThrGlnAspAlaMetValArgAspLeu 131
Db	133	GTCTTTGGCTCAGCTATAGCATATTTCATCACCTGTTCACTCGGATTTAACCAAGGAGTTG 192
QY	132	AsnLeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyGlyMetVal 151
Db	193	AATCTCTCTGTGGCAGAGTACTCGCTCTTCGGTTCGATTCTGACAATTGGAGCAATGATC 252
QY	152	GlyAlaIleAlaSerGlyGlnMetAlaGluTyrIleGlyArgLysGlySerLeuMetIle 171
Db	253	GGCGCAGCTATGAGTGGACGAATCGCCGATATGATAGGCAGACGAGCTACTATGGTTTC 312
QY	172	AlaAlaIleProAsnIleIleGlyTyrLeuAlaIleSerPheAlaLysAspAlaSerPhe 191
Db	313	TCGGAGATGTTTTCATTTCTCGGCTGGTTAGCAATCTACCTTTCTAAGGTTGCAATTTGG 372
QY	192	LeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTyrThrValPro 211
Db	373	CTTGACGTTGGGAGATTCTTGTTGGTTATGGAATGGCGGTTTTTTTCGTTTGGTTCCT 432
QY	212	ValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGln 231
Db	433	GTATACATAGCTGAAATCACTCCTAAGGGTCTTCGGTGGGTTTACAACGTTCATCAG 492
QY	232	LeuSerValThrPheGlyIlePheLeuAlaTyrLeuLeuGlyMetPheIleProTyrArg 251
Db	493	TTACTGATATGTTTGGGTGTCAGTCACATATCTCTGGGCTCTTTTATAGCTGGCGA 552
QY	252	LeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPhePheIle 271
Db	553	ATTTTAGCTTTGATCGGAATGATACCATGTGTTGTGACAAATGATGGGATTATTGTGCATC 612
QY	272	ProGluSerProArgTyrLeuAlaLysMetAsnLeuThrGluAspCysGluThrSerLeu 291
Db	613	CCAGAGTCTCCTAGATGGTTGGCAAAAGTTGGCAAAATGGGAAGAATTTGAGATTGCATTG 672
QY	292	GlnValLeuArgGlyPheGluThrAspIleThrThrGluValAsnAspIleLysArgAla 311
Db	673	CAACGATTACGTGGTGATCCGCAGATATTTCATATGAGTCTAATGAGATCAAAAGATTAT 732
QY	312	ValAlaSerSerSerLysArgThrThrIleSerPheGlnGluLeuAsnGlnLysLysTyr 331
Db	733	ACGCGAAGACTTACAGATTTATCAGAAGGTAGTATAGTAGATCTGTTCCAGCCACAGTAT 792
QY	332	ArgThrProLeuLeuGlyIleGlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsn 351
Db	793	GCTAAATCTCTAGTTGTGGAGTTGGTCTGTGATGGTTTGCACAATTTGGAGGGGTTAAT 852
QY	352	GlyValLeuPheTyrAlaSerSerIlePheLysAlaAlaGlyValThrAsnSerAspLeu 371
Db	853	GGAATTGCTTTTACGCTAGCTCTATTTTCGAATCTCTGGGTT---TCITCCAAAATT 909
QY	372	AlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThrGlyValThrThrTrpLeuLeu 391
Db	910	GGAATGATTGCGATGGTGGTCGTACAGATCCCTATGACAACTTTAGGTGTACTGTTGATG 969
QY	392	AspArgAlaGlyArgArgIleLeuLeuIleIleSerThrSerGlyMetThrLeuCysLeu 411
Db	970	GATAAATCAGGAAGACGACCATTAATTCTTAATTTCTGCTACTGGAACATGATCGGGTGT 1029
QY	412	LeuAlaValSerValPhePheLeuLysAspAsnIleSerGlnAspSerAsnSerTyr 431

Db

1030

TTCTCGTCGGTTTATCATTTTTCATTACAG-----

1059

QY

432

TyrIleLeuThrMetIleSerIleuValGlyIleValSerPheValIleThrPheSerPhe 451

Db

1060

-----GTGTACACCGGATCATTTTCATTA 1083

QY

452

GlyMetGlyAlaIleProTyrLeuMetMetSerGluIleLeuProValSerIleLysSer 471

Db

1084

GGGATGGTGGATTCTCTGGTTCATTATGTTCAGAGATATTCCAATAGACATAAAGGA 1143

QY

472

LeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeuThrSerPheAlaIleThrMetThr 491

Db

1144

TCAGCCGGAAGCCTCGTACTGTTGTAGCTGGTCCGATCATGGATTATATCTTTTACA 1203

QY

492

ThrAsnLeuMetLeuThrTrpSerValGlyGlyThrPheLeuSerTyrMetValValSer 511

Db

1204

TTTAACCTTCTAATGAATTGGAATCCGGCAGGAACGTTTATGTTTTCGCTACTGTTGT 1263

QY

512

AlaPheThrIleValPheValValLeuTyrValProGluThrLysGly**AsnSerArg 531

Db

1264

GGGGCTACTGTTATTTTGTAGCAAAACTCGTACCAGAAACCAAGGTGC-CACACITGA 1322

QY

532

GlyAspThrIlePheValSerLeuSerIleGlnArgGlnLeuGlnTrpLeu 548

Db

1323

GGAATCCAATATTC-----GATTGGTTA 1346

RESULT 12

ADM47939

ID

ADM47939

standard; DNA; 1768 BP.

XX

AC

ADM47939;

XX

DT

03-JUN-2004 (first entry)

XX

DE

Polynucleotide sequence #357 useful in producing transgenic plants.

XX

KW

Plant; maize; rice; sorghum; thale cress; soybean; wheat; transgenic; osmotic stress; sugar transport; cell cycle pathway; plant height; carbohydrate transport; crop productivity; plant growth; stress resistance; disease resistance; insect resistance; heat tolerance; nitrogen assimilation; water stress tolerance; photosynthetic carbon fixation; virus resistance; gene therapy; gene; ds.

XX

OS

Glycine max.

XX

PN

US2003233670-A1.

XX

PD

18-DEC-2003.

XX

PF

04-DEC-2002; 2002US-00310154.

XX

PR

04-DEC-2001; 2001US-0337358P.

XX

PA

(EDGE/) EDGERTON M D.

XX

PA

(CHOM/) CHOMET P S.

XX

PA

(LACC/) LACCETTI L B.

XX

PI

Edgerton MD, Chomet PS, Laccetti LB;

XX

XX

WPI; 2004-061374/06.

DR

P-PSDB; ADM48307.

XX

PT

New polynucleotide, useful for manipulating plant protein quality, improving plant growth, yield and crop productivity or grain composition or producing plants with improved properties.

XX

PS

Claim 1; SEQ ID NO 357; 144pp; English.

XX

CC

The present invention relates to polynucleotide sequences, and the proteins they encode. The sequences are isolated from a variety of organisms such as plants (e.g. maize, rice, sorghum, thale cress, soybean, and wheat), cyanobacteria, bacteria, yeast and other fungi. The polynucleotide and polypeptide sequences of the invention are useful in

CC the production of transgenic plants that have improved properties. Also
CC disclosed are methods of producing fertile transgenic plants, preferably
CC maize, with desired phenotypes. The polynucleotide and polypeptide
CC sequences are useful for improving plants by providing protection against
CC osmotic stress, improving altering sugar transport and/or metabolism,
CC modifying the cell cycle pathway, reducing plant height, modifying
CC carbohydrate transport, improving crop productivity, improving plant
CC growth and stress resistance, improving disease resistance, improving
CC insect resistance, improving cold or heat tolerance, improving nitrogen
CC assimilation, improving stalk strength, improving water stress tolerance,
CC improving photosynthetic carbon fixation, improving biotic and abiotic
CC stress resistance, improving resistance to oxidative stress, providing
CC increased vigour, reducing senescence, and conferring virus resistance.
CC The present sequence represents a polynucleotide sequence of the
CC invention. Note: The sequence data for this patent is not provided in the
CC printed specification but is obtained in electronic format from the USPTO
CC website at seqdata.uspto.gov.
XX
SQ Sequence 1768 BP; 516 A; 300 C; 384 G; 568 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 5.4e-71 Length: 1768
Score: 893.00 Matches: 191
Percent Similarity: 56.25% Conservative: 97
Best Local Similarity: 37.30% Mismatches: 173
Query Match: 32.13% Indels: 51
DB: 12 Gaps: 4

US-10-051-909-36 (1-553) x ADM47939 (1-1768)

QY	43	SerAsnArgGly-GlyAlaGlyAlaGlyGluSerGlySerAspHisAspGlyValle	62
Db	119	TCAACAGAGGATGCTGCTGCTTGTAAAGAAATGGATCAGAT	164
QY	62	uArgArgProLeuLeuAsnThrGlySerTrpTyrArgMetSerSerArgGlnSerSerPh	82
Db	165	-----AAGAGTGT	172
QY	82	eAlaProGlyThrSerSerMetAlaValLeuArgGluSerHisValSerAlaPheLeuCy	102
Db	173	GAATAATGGATCCATTTGAATGGTTTG-----CTTAG	205
QY	102	sThrLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGlyPheSerSerPr	122
Db	206	CACACTTGTGTGTTGTGTTCTTTCACATTGGAACCTTGTGTGGCTATTTCAGCACC	265
QY	122	oThrGlnAspAlaMetValArgAspLeuAsnLeuSerIleSerGluPheSerAlaPheGl	142
Db	266	CACCTAAGCGGCTATCAGGCGCAGATCTTAATCTCTCTTGTGAGTTTCCATGTTTG	325
QY	142	ySerLeuSerAsnValGlyGlyMetValGlyAlaIleAlaSerGlyGlnMetAlaGluTy	162
Db	326	TTCAATTAGTGACCATGGTGCAATGCTTGGAGCTATAACACGCGGAGGATTACAGATT	385
QY	162	rIleGlyArgGlySerLeuMetIleAlaAlaIleProAsnIleIleGlyTrpLeuAl	182
Db	386	CATTGGCCGCAAGGGCAATGAGGATTTCAACAGGATTTTGCATTACAGATGGATAGC	445
QY	182	aIleSerPheAlaLysAspAlaSerPheLeuTyrMetGlyArgLeuLeuGluGlyPheGl	202
Db	446	TGTCTTCTTCTAAGGTTCTTACTCCCTTGACTTCGGAAGATTTTTCACAGTTACGG	505
QY	202	yValGlyIleIleSerTyrThrValProValTyrIleAlaGluIleSerProGlnAsnMe	222
Db	506	CATTGGAGTTATCTCATACGTTCTCTGTATATATAGCAGAAATAGCACCCAAATCT	565
QY	222	tArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIlePheLeuAlaTy	242
Db	566	TCGAGGAGGACTAGCAACAACAATCAGCTTTGATTGTTACTGGAGGTCAGTCTCATT	625
QY	242	rLeuLeuGlyMetPheIleProTrpArgLeuLeuAlaValIleGlyAlaLeuProCysTh	262
Db	626	CTTATTAGGAAGTGTATAAATTGGAGAGAACTTGATTAGCGGGGCTAGTGCCTTGCAT	685

QY	262	rMetLeuIleProGlyLeuPhePheIleProGluSerProArgTrpLeuAlaLysMetAs	282
Db	686	TTGCTTGTGTTGGTTTGTGCTTTATCCCTGAGTCCCCAGATGGCTGGTAAGTTGG	745
QY	282	nLeuThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPheGluThrAspIleTh	302
Db	746	CCGTGAAAAGAAATTTCAACTAGCTTTAAGTAGACTTCGGGGTAAAGATGCTGATATT	805
QY	302	rThrGluValAsnAspIleLysArgAlaValAlaSerSerLysArgThrThrIleSe	322
Db	806	TGATGAAGCTGCTGAAATTTCTGGATTATATGAAACTCTTCAAAGTCTTCTTAAGACTAA	865
QY	322	rPheGlnGluLeuAsnGlnLysLysTyrArgThrProLeuLeuGlyIleGlyLeuLe	342
Db	866	GCTGTTGGATTGTTCCAAAGCAAAATATGTCACCTCTGTAGTTATTGGGTCGGCTTAAT	925
QY	342	uValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSerIlePheLy	362
Db	926	GGCATGTCAACAATCTGTTGGAATTAATGGCATAGGATTCTACACAGCTGAGATTTTGT	985
QY	362	sAlaAlaGlyValThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIleGlnValle	382
Db	986	AGCAGCTGGACTTCTTTCAGGAAAAGCTGGTACCATAGCATATGCTGTATACAGATTCC	1045
QY	382	uAlaThrGlyValThrThrTrpLeuLeuAspArgAlaGlyArgArgIleLeuLeuIleI	402
Db	1046	ATTTACATTATTGGGAGCCATTTTTGATGGACAAGTCTGGAAGAAGACCTCTTGTAAATGGT	1105
QY	402	eSerThrSerGlyMetThrLeuCysLeuLeuAlaValSerValValPhePheLeuLysAs	422
Db	1106	TTCCGCGAGCTGGGACATTCTTAGGTTGCTTGTGCTGCTTGTGCTTCTTCTCAAGGA	1165
QY	422	pAsnIleSerGlnAspSerAsnSerTyrTyrIleLeuThrMetIleSerLeuValGlyI	442
Db	1166	C-----CAAAGCTTATTGCTGAGTGGGTACCTATATTAGCATTTGTGCGCGT	1213
QY	442	eValSerPheValIleThrPheSerPheGlyMetGlyAlaIleProTrpLeuMetMetSe	462
Db	1214	TCTGATCTACATAGCAGCATTTTCAATTGGGCTCGGATCAGTTCATGGGTGATAATGTC	1273
QY	462	rGluIleLeuProValSerIleLysSerLeuGlyGlySerIleAlaThrLeuAlaAsnTr	482
Db	1274	TGAGATATTTCCTCATACATTGAAGGGGACTGCTGGAAGCTTGGTGGTTTGGTGGCTTG	1333
QY	482	pLeuThrSerPheAlaIleThrMetThrThrAsnLeuMetLeuThrTrpSerValGlyGl	502
Db	1334	GCTAGGAGCTTGGGTAGTTTCATATACTTTCAACTTTCTTATGAGCTGGAGTTCTCCTGG	1393
QY	502	yThrPheLeuSerTyrMetValValSerAlaPheThrIleValPheValValLeuTrpVa	522
Db	1394	TACATTGTTTTTGTATGCTGGATGTTCCCTCTTAACCTATCTATTCTAGCAAAATAGT	1453
QY	522	lProGluThrLysGly***AsnSerArgGlyAspThrIlePheValSerLeuSerIleGl	542
Db	1454	CCAGAAACCAAGGA-----	1469
QY	542	nArgGlnLeuGlnTrpLeuProGluCysLeuSer	553
Db	1470	-AAAACCTTGGAGAGATCCAGGCTTGTATTAGT	1502
RESULT 13			
ABZ13486			
ID	ABZ13486	standard; DNA; 1407 BP.	
XX	AC	ABZ13486;	
XX	DT	21-JAN-2003 (first entry)	
XX	DE	Arabidopsis thaliana stress regulated gene SEQ ID NO 1291.	
XX	KW	Arabidopsis thaliana; plant; gene; stress; transgenic; ds.	

OS Arabidopsis thaliana.
XX WO200216655-A2.
XX
PD 28-FEB-2002.
XX
XX 24-AUG-2001; 2001WO-US026685.
XX
PR 24-AUG-2000; 2000US-0227866P.
PR 26-JAN-2001; 2001US-0264647P.
PR 22-JUN-2001; 2001US-0300111P.
XX
PA (SCRI) SCRIPPS RES INST.
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
PI Harper JF, Kreps J, Wang X, Zhu T;
XX WPI; 2002-304127/34.
DR
XX
PT Identifying a stress condition to which a plant cell has been exposed and
PT producing plants with increased tolerance to these abiotic stresses.
XX
PS Claim 144; SEQ ID NO 1291; 577pp + Sequence Listing; English.
XX
CC The invention relates to identifying a stress condition to which a plant
CC cell has been exposed, comprising: (a) contacting nucleic acid
CC representative of expressed polynucleotides in the plant cell with an
CC array or probes representative of the plant cell genome; and (b)
CC detecting a profile of expressed polynucleotides in the plant cell
CC characteristic of a stress response. The method is useful in the
CC production of transgenic plants, cells and seeds and in producing plants
CC with increased tolerance to abiotic stress. The present sequence is that
CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
CC in methods of the invention. Note: The sequence data for this patent is
CC not represented in the printed specification but is based on sequence
CC information supplied to Derwent by the European Patent Office
XX
SQ Sequence 1407 BP; 350 A; 261 C; 331 G; 465 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 9.92e-70 Length: 1407
Score: 877.50 Matches: 179
Percent Similarity: 59.06% Conservative: 98
Best Local Similarity: 38.17% Mismatches: 166
Query Match: 31.58% Indels: 26
DB: 6 Gaps: 5

US-10-051-909-36 (1-553) x ABZ13486 (1-1407)

QY 66 LeuLeuAsnThrGlySerTrpTyrArg-MetSerSerArgGlnSerSerPheAlaProGl 85
Db 12 CTCAAGAAACTTGGAGCGTGTCTGTCTACTGAGGAAGAACCAACGAC----- 60

QY 85 yThrSerSerMetAlaValLeuArgGluSerHisValSerAla-----PheLeuCysTh 103
Db 61 -----ATCAACGAATGTCGTATCACTGCTGTGTGTAATTTTCAGTAC 101

QY 103 rLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGlyPheSerSerProTh 123
Db 102 TTTCGTTTCTGTTTGTGGCTCTTCTGCTTCGGTGTGCGGAGGTTTATCATCAGTTGC 161

QY 123 rGlnAspAlaMetValArgAspLeuAsnLeuSerIleSerGluPheSerAlapheGlySe 143
Db 162 TCAACAGGGATCATAAATGATTAGGTCCTCTCTGTGCAATACTCCATGTTTGGTTC 221

QY 143 rLeuSerAsnValGlyGlyMetValGlyAlaIleAlaSerGlyGlnMetAlaGluTyrIl 163
Db 222 AATCATGACTTTTGGAGGAATGATTGGTGCCATCTTCAGCGGAAAGTGCAGATCTCAT 281

QY 163 eGlyArgLysGlySerLeuMetIleAlaAlaIleProAsnIleIleGlyTrpLeuAlaIl 183
Db 282 GGGTCGAAAAGGAGTATGTGGTTTGCTCAAATTTTCTGCATCTTCCGGTTGGGTGCAGT 341

QY 183 eSerPheAlaLysAspAlaSerPheLeuTyrMetGlyArgLeuLeuGluGlyPheGlyVa 203
Db 342 AGCATTAGCAAAAGACTCCATGTGGCTTGATATTGGAAGACTATCCACAGGATTTCAGT 401

QY 203 lGlyIleIleSerTyrThrValProValTyrIleAlaGluIleSerProGlnAsnMetAr 223
Db 402 TGGTTTATTAAAGCTATGTGATACCAGTTTACATTGCAGAAATAACACCAACATGTCG 461

QY 223 gGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIlePheLeuAlaTyrLe 243
Db 462 AGGACCGTTTGTATTGCTAATCAGCTGATGCAGAGTTGTGGATTGTCTTTATTCTACGT 521

QY 243 uLeuGlyMetPheIleProTrpArgLeuLeuAlaValIleGlyAlaLeuProCysThrMe 263
Db 522 CATGGAAATTTTGTTCATTGGCGTAACTTGGCCCTTAATCGGTCTCATTCATGTGCGTT 581

QY 263 tLeuIleProGlyLeuPhePheIleProGluSerProArgTrpLeuAlaLysMetAsnLe 283
Db 582 GCAAGTTGTGACTTTGTTCTTTATTCCAGAGTCCCTTAGACTACTGGGAAATGGGACACA 641

QY 283 uThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPheGluThrAspIleThrTh 303
Db 642 TGA AAAAGAAATGTAGAGCTTCATTGCAAGTCTTCGCGGAGATGATGCAGATATCTCTGA 701

QY 303 rGluValAsnAspIleLysArgAlaValAlaSerSerSerLysArgThrThrIleSerPh 323
Db 702 AGAAGCCCAACACTATCAAGAAACCATGATCTTGTTCATGAAGGACCAAAATCGCGGT 761

QY 323 eGlnGluLeuAsnGlnLysLysTyrArgThrProLeuLeuLeuGlyIleGlyLeuLeuVa 343
Db 762 TATGGATTGTTTTCAGAGAAATATGCTCCATCTGTGTTTATTTGGTGTGGGACTAATGCT 821

QY 343 lLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSerIlePheLysAl 363
Db 822 TCTACAAACAACTCTCTGGAAGCTCAGGACTTATGTACTATGTTCGGTAGCGTATTGTATAA 881

QY 363 aAlaGlyValThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIleGlnValLeuAl 383
Db 882 AGGAGGGTTTCCAAGCAGC---ATTGGCTCAATGATTCTTGCAGTGATCATGACCAAA 938

QY 383 aThrGlyValThrTrpLeuLeuAspArgAlaGlyArgArgIleLeuLeuIleIleSe 403
Db 939 AGCTCTATTGGGTCTGATTTTGGTTGAGAAATGGGAGGAGGACCACTTCTATTGGGCTC 998

QY 403 rThrSerGlyMetThrLeuCysLeuLeuAlaValSerValPhePheLeuLysAspAs 423
Db 999 TACCGTGAATGTGCTTTTTCAGCTTGTCTCCTCAGTTTTTCTTCTGCTTTCGG---- 1053

QY 423 nIleSerGlnAspSerAsnSerTyrTyrIleLeu-----ThrMetIleSerLe 439
Db 1054 -----TCATATGGCATGCTTGTATGAGCTCACTCCGATTTTTCACATG 1094

QY 439 uValGlyIleValSerPheValIleThrPheSerPheGlyMetGlyAlaIleProTrpLe 459
Db 1095 TATCGGTGTAGTGGGTTTCATCTCTTTCATTTGCCGTAGGCATGGGAGGCTTACCATGGAT 1154

QY 459 uMetMetSerGluIleLeuProValSerIleLysSerLeuGlyGlySerIleAlaThrLe 479
Db 1155 CATCATGCTGAGATATTCCCAATGAATGTTAAAGTTTCTGTGGGACTCTGGTTACCTT 1214

QY 479 uAlaAsnTrpLeuThrSerPheAlaIleThrMetThrThrAsnLeuMetLeuThrTrpSe 499
Db 1215 AGCCAACTGGTCCTTTTGGTTGGATTGTTGCTTTTCGCCTACAACTTCATGTAGTGGAA 1274

QY 499 rValGlyGlyThrPheLeuSerTyrMetValValSerAlaPheThrIleValPheValVa 519
Db 1275 CGCATCAGGAACGTTCTTGATCTCTTTACTATATGTGTCGGGTATAGTCTTTATTATA 1334

QY 519 lLeuTrpValProGluThrLysGly 527
Db 1335 TCGATGGTACCAGGAACTAAAGGA 1359

ADA67841
ID ADA67841 standard; DNA; 1407 BP.
XX
AC ADA67841;
XX
DT 20-NOV-2003 (first entry)
XX
DE Arabidopsis thaliana gene, SEQ ID 79.
XX
KW Plant; bacterial infection; fungal infection; viral infection; ds.
XX
OS Arabidopsis thaliana.
XX
PN WO2003000898-A1.
XX
PD 03-JAN-2003.
XX
PF 22-JUN-2001; 2001WO-IB001105.
XX
PR 22-JUN-2001; 2001WO-IB001105.
XX
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
PI
DR WPI; 2003-175290/17.
XX
PT Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.
XX
PS Claim 6; SEQ ID NO 79; 899pp; English.
XX
CC The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance to resistance or tolerance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.
XX
SQ Sequence 1407 BP; 350 A; 261 C; 331 G; 465 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 9,92e-70 Length: 1407
Score: 877.50 Matches: 179
Percent Similarity: 59.06% Conservative: 98
Best Local Similarity: 38.17% Mismatches: 166
Query Match: 31.58% Indels: 26
DB: 8 Gaps: 5

US-10-051-909-36 (1-553) x ADA67841 (1-1407)

QY 66 LeuLeuAsnThrGlySerTrpTyrArg-MetSerSerArgGlnSerSerPheAlaProG1 85
Db 12 CTCAGAAACCTTGAAGCTGGTTTGTCTACTGAGGAAGAACCAACGAC----- 60

QY 85 yThrSerSerMetAlaValLeuArgGluSerHisValSerAla-----PheLeuCysTh 103
Db 61 -----ATCAACGAATGTCGTATCAGTCTGTGTACTTTTCAGTAC 101

QY 103 rLeuileValAlaLeuGlyProileGlnPheGlyPheThrSerGlyPheSerSerProTh 123
Db 102 TTTTCGTTTCTGTTGTGGCTCTTTTCTGCTCGGTGTGGCGGAGGTTATTCATCAGTTGC 161

QY 123 rGlnAspAlaMetValArgAspLeuAsnLeuSerIleSerGluPheSerAlaPheGlySe 143
Db 162 TCAAACAGGGATCAATAATGATTAGGTTCTCTGTTGCACAACTACTCCATGTTTGGTTC 221

QY 143 rLeuSerAsnValGlyGlyMetValGlyAlaIleAlaSerGlyGlnMetAlaGluTyril 163
Db 222 AATCATGACTTTTGGAGGAATGATTGGTGCCATCTTCAGCGGAAAGTTGCAGATCTCAT 281

QY 163 eGlyArgLysGlySerLeuMetIleAlaAlaIleProAsnIleIleGlyTrpLeuAlaI 183
Db 282 GGGTCGAAAAGGACTATGTGGTTTGTCTCAAAATTTTCTGCATCTTCGGTTGGTTGCAGT 341

QY 183 eSerPheAlaLysAspAlaSerPheLeuTyrMetGlyArgLeuLeuGluGlyPheGlyVa 203
Db 342 AGCATTAGCAAAAGACTCCATGTGGCTTGATATTGGAAGACTATCCACAGGATTTCAGT 401

QY 203 lGlyIleIleSerTyrThrValProValTyrIleAlaGluIleSerProGlnAsnMetAr 223
Db 402 TGGTTTATTAGCTATGTGATACCAGTTTACATTCAGAAATAACACCAACATGTTTCG 461

QY 223 gGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIlePheLeuAlaTyrLe 243
Db 462 AGGAGCGTTTGTATTTGCTAATCAGCTGATGCAGAGTTGTGGATTGTCTTTATTCTACGT 521

QY 243 uLeuGlyMetPheIleProTrpArgLeuLeuAlaValIleGlyAlaLeuProCysThrMe 263
Db 522 CATTTGGAATTTTGTTCATTGGCGTAACTTGGCCCTTAATCGGTCTCATTCATTCGCGTT 581

QY 263 tLeuIleProGlyLeuPhePheIleProGluSerProArgTrpLeuAlaLysMetAsnLe 283
Db 582 GCAAGTTTGACTTTGTTCTTTTATTCAGAGTCCCCTAGACTACTGGGAAAATGGGGACA 641

QY 283 uThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPheGluThrAspIleThrTh 303
Db 642 TGAATAAGAAATGTAGAGCTTCATTGCAAGTCTTTCGGGAGATCATGCAGATATCTCTGA 701

QY 303 rGluValAsnAspIleLysArgAlaValAlaSerSerSerLysArgThrThrIleSerPh 323
Db 702 AGAAGCCCAACTATCAAAGAAACCAATCATCTTGTGTTGATGAGGACCAAAATCGCGGT 761

QY 323 eGlnGluLeuAsnGlnLysLysTyrArgThrProLeuLeuLeuGlyIleGlyLeuLeuVa 343
Db 762 TATGGATTTGTTTCAGAGAAGATATGCTCCATCTGTTGTTATTTGTTGGGACTAATGCT 821

QY 343 lLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSerIlePheLysAl 363
Db 822 TCTACAACAACCTCTCTGGAAGCTCAGGACTTATGTACTATGTCGGTAGCGTATTGTATA 881

QY 363 aAlaGlyValThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIleGlnValLeuAl 383
Db 882 AGGAGGGTTTCCAAGCAGC--ATTGGCTCAATGATTCTTGCAGTGATCATGATACCAA 938

QY 383 aThrGlyValThrTrpLeuLeuAspArgAlaGlyArgArgIleLeuLeuIleIleSe 403
Db 939 AGCTCTATTGGGCTGTGATTTTGGTTGAGAAAATGGGACGAGACCACTTCTATTGGCCTC 998

QY 403 rThrSerGlyMetThrLeuCysLeuLeuAlaValSerValValPhePheLeuLysAspAs 423
Db 999 TACCGGTGAATGTGCTTTTTCAGCTTGCTCCTCAGTTTTCCTTCTGCTTTCGG----- 1053

QY 423 nIleSerGlnAspSerAsnSerTyrTyrIleLeu-----ThrMetIleSerLe 439
Db 1054 -----TCATATGGCATGCTTGATGAGCTCACTCCGATTTTCACATG 1094

QY 439 uValGlyIleValSerPheValIleThrPheSerPheGlyMetGlyAlaIleProTrpLe 459
Db 1095 TATCGGTGTAGTGGTTTTCATCTCTTCATTTGCCGTAGGCATGGGAGGCTTACCATGGAT 1154

QY 459 uMetMetSerGluIleLeuProValSerIleLysSerLeuGlyGlySerIleAlaThrLe 479
Db 1155 CATCATGTCTGAGATATTCCCAATGATGTTAAAGTTTCTGCTGGGACTCTGGTTACCTT 1214

QY 479 uAlaAsnTrpLeuThrSerPheAlaIleThrMetThrThrAsnLeuMetLeuThrTrpSe 499
Db 1215 AGCCAACACTGGTCTTGGTTGGATTGTTGCTTTCGCCTACAACACTCATGCTAGAGTGAA 1274

QY 499 rValGlyGlyThrPheLeuSerTyrMetValValSerAlaPheThrIleValPheValva 519
; ||||| ||| : : : ||||| : : :
Db 1275 CGCATCAGGAACGTCTTGATCTTCTTTACTATATGTTGGTGGGTATAGTCTTTATTTA 1334

QY 519 lLeuTrpValProGluThrLysGly 527
||| ||||| ||||| |||||
Db 1335 TCGGATGGTACCCGAAACTAAAGGA 1359

RESULT 15
AAC46009
ID AAC46009 standard; DNA; 1642 BP.
XX
AC AAC46009;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 48579.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
FN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX

PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.

PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.

PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:
Pred. No.: 1,24e-69
Score: 877.50
Percent Similarity: 59.06%
Best Local Similarity: 38.17%
Query Match: 31.58%
DB: 3

Length: 1642
Matches: 179
Conservative: 98
Mismatch: 166
Indels: 26
Gaps: 5

US-10-051-909-36 (1-553) x AAC46009 (1-1642)

QY 66 LeuLeuAsnThrGlySerTyrArg-MetSerSerArgGlnSerSerPheAlaProGl 85
Db 18 CTCAAGAAACTTGAAGCTGGTTTGCTACTGAGGAAGAACCAAAACGAC----- 66
QY 85 yThrSerSerMetAlaValLeuArgGluSerHisValSerAla-----PheLeuCysTh 103
Db 67 -----ATCAACGAATGTCGTATCACTGCTGTGTACTTTTCAGTAC 107
QY 103 rLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGlyPheSerSerProTh 123
Db 108 TTTCTGTTCTGTTTGTGGCTCTTTCTGCTTCGGTTGTGCGGAGGTTATTTCATCAGTTGC 167
QY 123 rGlnAspAlaMetValArgAspLeuAsnLeuSerIleSerGluPheSerAlaPheGlySe 143
Db 168 TCAAACAGGATCATAAATGATTAGGTCTCTCTGTTGCACAATACTCCATGTTTGGTTC 227
QY 143 rLeuSerAsnValGlyGlyMetValGlyAlaIleAlaSerGlyGlnMetAlaGluTyrIl 163
Db 228 AATCATGACTTTTGGAGGAATGATTGGTGCCATCTTCAGCGGGAAGTTGCAGATCTCAT 287
QY 163 eGlyArgLysGlySerLeuMetIleAlaAlaIleProAsnIleIleGlyTrpLeuAlaIl 183
Db 288 GGGTCGAAAAGGGACTATGTGGTTTCTCTCAAAATTTTCTGCACTCTTCGGTTGGTTGCACT 347
QY 183 eSerPheAlaLysAspAlaSerPheLeuTyrMetGlyArgLeuLeuGluGlyPheGlyVa 203
Db 348 AGCATTAGCAAAAGACTCCATGTGGCTTGATATTGGAAGACTATCCACAGGATTTGCAGT 407
QY 203 lGlyIleIleSerTyrThrValProValTyrIleAlaGluIleSerProGlnAsnMetAr 223
Db 408 TGGTTTATTAAAGCTATGTGATACCACTTACATTGCAGAAATAACCAAAACATGTTTCG 467
QY 223 gGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIlePheLeuAlaTyrIe 243
Db 468 AGGAGCGTTTGTATTGTGTAATCAGCTGATGCAGAGTTGTGGATTGTTTATTCTACGT 527
QY 243 uLeuGlyMetPheIleProTyrArgLeuLeuAlaValIleGlyAlaLeuProCysThrMe 263
Db 528 CATTTGGAATTTTGTTCATTGGCGTAACTTGGCCTTAATCGGCTCATTCATGTCGCTT 587
QY 263 tLeuIleProGlyLeuPhePheIleProGluSerProArgTyrPheLeuAlaLysMetAsnle 283
Db 588 GCAAGTTGTGACTTTGTTCTTTTATCCAGAGTCCCTAGACTACTGGGAAAATGGGGACA 647
QY 283 uThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPheGluThrAspIleThrTh 303
Db 648 TGAAAAGAAATGTAGAGCTTTCATTGCAAAAGTCTTCGCGGAGATGATGAGATATCTCTGA 707
QY 303 rGluValAsnAspIleLysArgAlaValAlaSerSerSerLysArgThrThrIleSerPh 323
Db 708 AGAAGCCACACTATCAAGAAACCATGATCTTGTGTTGATGAGGACCAAAATCGCGGT 767
QY 323 eGlnGluLeuAsnGlnLysLysTyrArgThrProLeuLeuLeuGlyIleGlyLeuLeuVa 343
Db 768 TATGGATTTGTTTCAGAGAAGATATGCTCCATCTGTTGTTTATGTTGTTGTTGTTGTTG 827
QY 343 lLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSerIlePheLysAl 363
Db 828 TCTACAAACAACCTCTCTGGAAGCTCAGGACTTATGTACTATGTCGTTAGCGTATTGTATAA 887
QY 363 aAlaGlyValThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIleGlnValLeuAl 383
Db 888 AGGAGGGTTTCCAAAGCAGC---ATTGGCTCAATGATTTCTTGCAGTGATCATGATACCAA 944
QY 383 aThrGlyValThrThrTrpLeuLeuAspArgAlaGlyArgArgIleLeuLeuIleSe 403
Db 945 AGCTCTATTGGGTCTGATTTTGGTTGAGAAAATGGGACGAGACCACTTCTATTGGCCTC 1004
QY 403 rThrSerGlyMetThrLeuCysLeuLeuAlaValSerValPhePheLeuLysAspAs 423
Db 1005 TACCGGTGGAATGTGCTTTTTCAGCTTGCTCTCTCAGTTTCTTCTCTCTGCTTCGG----- 1059

QY	423	nIleSerGlnAspSerAsnSerTyrTyrIleLeu-----ThrMetIleSerIle	439
Db	1060	-----TCATATGGCATGCTTGATGAGCTCACTCCGATTTCACATG	1100
QY	439	uValGlyIleValSerPheValIleThrPheSerPheGlyMetGlyAlaIleProTrpLe	459
Db	1101	TATCGGTGTAGTGGTTTCATCTCTTCATTGCCGTAGGCATGGGAGGCTTACCATGGAT	1160
QY	459	uMetMetSerGluIleLeuProValSerIleLysSerLeuGlyGlySerIleAlaThrLe	479
Db	1161	CATCATGTCTGAGATATCCCAATGAATGTAAAGTTTCTGCTGGGACTCTGGTTACCTT	1220
QY	479	uAlaAsnTrpLeuThrSerPheAlaIleThrMetThrThrAsnLeuMetLeuThrTrpSe	499
Db	1221	AGCCAACTGGTCCCTTTGGTTGGATTGTTGCTTACCACTTCATGCTAGAGTGGAA	1280
QY	499	rValGlyGlyThrPheLeuSerTyrMetValValSerAlaPheThrIleValPheValVa	519
Db	1281	CGCATCAGGAACGTTCTTGATCTTCTTTACTATATGTGGTGGGGTATAGTCTTTATTTA	1340
QY	519	LleuTrpValProGluThrLysGly	527
Db	1341	TGGGATGGTACCCGAAACTAAAGGA	1365

Search completed: October 13, 2004, 12:24:30
Job time : 717.434 secs

is Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 13, 2004, 11:54:50 ; Search time 116.894 Seconds
(without alignments)
3362.578 Million cell updates/sec

Title: US-10-051-909-36
Perfect score: 2779
Sequence: 1 PSSSSSFRPAGKKKKKNQG.....TIFVLSIQRQLWLPECLS 553

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 824507 seqs, 355394441 residues
Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO spool_h/US10051909/runat_13102004_123337_19931/app query.fasta_1.1678
-DB=Issued Patents_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10051909@cgn_1_116@runat_13102004_123337_19931 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/6C_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	578	20.8	5228	4 US-09-919-039-216	Sequence 216, App
2	560.5	20.2	5227	4 US-09-919-172-79	Sequence 79, Appl
3	557.5	20.1	2089	3 US-09-291-922-25	Sequence 25, Appl
4	557	20.0	2017	3 US-09-291-922-21	Sequence 21, Appl
5	551.5	19.8	3915	4 US-09-023-655-1104	Sequence 1104, Ap
6	542.5	19.5	1853	3 US-09-291-922-23	Sequence 23, Appl
7	528	19.0	2592	3 US-09-591-025-8	Sequence 8, Appli
8	527.5	19.0	1914	3 US-09-291-922-19	Sequence 19, Appl
9	524	18.9	2592	4 US-09-894-927B-8	Sequence 8, Appli
10	523.5	18.8	2856	4 US-09-643-597-135	Sequence 135, App
11	523.5	18.8	2856	4 US-09-480-884A-135	Sequence 135, App
12	523.5	18.8	2856	4 US-09-542-615A-135	Sequence 135, App

13	523.5	18.8	2856	4 US-09-606-421B-135	Sequence 135, App
14	523.5	18.8	2856	4 US-09-221-107-135	Sequence 135, App
15	523.5	18.8	2856	4 US-09-466-396A-135	Sequence 135, App
16	523.5	18.8	2856	4 US-09-476-496A-135	Sequence 135, App
17	523.5	18.8	2856	4 US-09-630-940B-135	Sequence 135, App
18	522	18.8	1872	3 US-09-291-922-27	Sequence 27, Appl
19	519	18.7	3000	2 US-08-928-692-9	Sequence 9, Appli
20	519	18.7	3000	3 US-09-339-972-9	Sequence 9, Appli
21	514.5	18.5	1431	4 US-09-489-039A-4762	Sequence 4762, Ap
22	497	17.9	1545	4 US-09-489-039A-4731	Sequence 4731, Ap
23	470.5	16.9	1752	4 US-09-679-686B-17	Sequence 17, Appl
24	469	16.9	1776	4 US-09-679-686B-11	Sequence 11, Appl
25	455	16.4	1960	4 US-09-679-686B-15	Sequence 15, Appl
26	451	16.2	1675	4 US-09-679-686B-1	Sequence 1, Appli
27	450	16.2	1704	4 US-09-614-221A-96	Sequence 96, Appl
28	447.5	16.1	1668	4 US-09-614-221A-420	Sequence 420, App
29	446.5	16.1	1704	4 US-09-614-221A-534	Sequence 534, App
30	446	16.0	1506	4 US-09-489-039A-4560	Sequence 4560, Ap
31	442	15.9	1659	4 US-09-248-796A-6610	Sequence 6610, Ap
32	439.5	15.8	987	4 US-09-710-279-1941	Sequence 1941, Ap
33	439.5	15.8	3065	4 US-09-710-279-3697	Sequence 3697, Ap
34	439.5	15.8	4411529	3 US-09-103-840A-1	Sequence 1, Appli
35	435	15.7	1566	4 US-09-489-039A-2378	Sequence 2378, Ap
36	421.5	15.2	1626	4 US-09-614-221A-521	Sequence 521, App
37	420	15.1	1515	4 US-09-248-796A-6607	Sequence 6607, Ap
38	418	15.0	1714	4 US-09-774-528-437	Sequence 437, App
39	417	15.0	1943	4 US-09-774-528-168	Sequence 168, App
40	416.5	15.0	1689	4 US-10-162-012-45	Sequence 45, Appl
41	416.5	15.0	4385	4 US-10-162-012-43	Sequence 43, Appl
42	413.5	14.9	1521	4 US-09-248-796A-5972	Sequence 5972, Ap
43	409.5	14.7	2601	3 US-09-291-922-7	Sequence 7, Appli
44	407.5	14.7	4403765	3 US-09-103-840A-2	Sequence 2, Appli
45	400	14.4	1695	4 US-09-614-221A-443	Sequence 443, App

ALIGNMENTS

RESULT 1

US-09-919-039-216
; Sequence 216, Application US/09919039
; Patent No. 6727066
; GENERAL INFORMATION:
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
; FILE REFERENCE: PA-0035 US
; CURRENT APPLICATION NUMBER: US/09/919,039
; CURRENT FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 60/222,113
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 401
; SOFTWARE: PERL Program
; SEQ ID NO 216
; LENGTH: 5228
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6727066 410785.1
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 4929, 4935, 4940, 4945, 4974, 4993
; OTHER INFORMATION: a, t, c, g, or other
US-09-919-039-216

Alignment Scores:
Pred. No.: 1.88e-52
Score: 578.00
Percent Similarity: 49.22%
Best Local Similarity: 29.57%
Query Match: 20.80%
DB: 4
Length: 5228
Matches: 152
Conservative: 101
Mismatches: 167
Indels: 94
Gaps: 15

US-10-051-909-36 (1-553) x US-09-919-039-216 (1-5228)

QY 86 ThrSerSerMetAlaValLeuArgGluSerHisValSerAlaPheLeuCys---ThrLeu 104
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 30 ACTAGTACAATGACA-----GAAGATAAGGTCACCTGGGACCCCTGGTTTTCACCTGTC 80

QY 105 IleValAla---LeuGlyProIleGlnPheCysPheThrSerGlyPheSerProThr 123
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 81 ATCACTGCTGCTGGGTCCTTCCAGTTTGGATATGACATTGGTGTGATCAATGCACCT 140

QY 124 GlnAspAlaMetValArgAspLeu----- 131
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 141 CAACAGGTAATAATATCTCACTATAGACATGTTTGGGTGTTCCACTGGATGCCGAAAA 200

QY 131 ----- 131
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 201 GCTATCAACAACCTATGTTATCAACAGTACAGATGAAGTCCCAACAATCTCATCTCAATG 260

QY 131 ----- 131
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 261 AACCCAAAACCAACCCCTTGGGCTGAGGAAGAGACTGTGGCAGCTGCTCAACTAATCACC 320

QY 132 -----AsnLeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGly 148
: : : : : : : : : : | | | | | : : : : : | | | | | : : : : :
Db 321 ATGCTCTGGTCCCTGTCTGTATCCAGCTTGCA-----GTTGGT 359

QY 149 GlyMetValGlyAlaIleAlaSerGlyGlnMetAlaGluTyrIleGlyArgLysGlySer 168
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 360 GGAATGACTGCATCATCTTTGGTGGTGGCTGGGACACACTTGGGAAGAAATCAAAGCC 419

QY 169 LeuMetIleAlaAlaIleProAsnIleIleGlyTyrLeuAlaIleSerPheAlaLys--- 187
: : : : : : : : : : | | | | | : : : : : | | | | | : : : : :
Db 420 ATGTTAGTAGCAAAACATCTGTCAATAGTGGAGCTCTCTTGATGGGGTTCACAAATG 479

QY 188 -----AspAlaSerPheLeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIle 205
: : : : : : : : : : | | | | | : : : : : | | | | | : : : : :
Db 480 GGACCATCTCATATACTTATAATTGCTGGAAGAAGCATATCAGGACTATATTGTGGGCTA 539

QY 206 IleSerTyrThrValProValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAla 225
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 540 ATTTGAGGCTGCTGCTATGATATCGGTGAAATTGCTCCAAACCGCTCTCAGGGGAGCA 599

QY 226 LeuGlySerValAsnGlnLeuSerValThrPheGlyIlePhe----- 239
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 600 CTGGGCACTTTTCATCAGCTGGCCATCGTCACGGGCAATCTTATTAGTCAGATTATTGGT 659

QY 240 LeuAlaTyrLeuLeuGlyMetPheIleProTyrArgLeuLeuAlaValIleGlyAlaLeu 259
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 660 CTGGAATTTATCTGGGCAATTATGATCTGTGGCACATCTTGTGGCTGTCTGTGGTGTG 719

QY 260 ProCysThrMetLeuIleProGlyLeuPhePheIleProGluSerProArgTyrLeu--- 278
: : : : : : : : : : | | | | | : : : : : | | | | | : : : : :
Db 720 CGAGCCATCCTCTCAGTCTCTGCTACTCTTTTCTGCCAGAAAGCCCGATACCTTTAC 779

QY 279 AlaLysMetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPheGlu 298
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 780 ATCAAGTTAGATGAGGAAGTCAAGCAAAACAAAGCTTGAAAGACTCAGAGGATATGAT 839

QY 299 ThrAspIleThrThrGluValAsnAspIleLys-----ArgAlaValAlaSerSerSer 316
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 840 ---GATGTACCAAAAGATATTAATGAAATGAGAAAAGAAAGAGAGCATCGAGTGAG 896

QY 317 LysArgThrThrIleSerPheGlnGluLeuAsnGlnLysTyrArgThrProLeuLeu 336
: : : : : : : : : : | | | | | : : : : : | | | | | : : : : :
Db 897 CAGAAAGTCTCTATA---ATTGAGCTCTTCCCAATTCAGTACCGACAGCCTATTCTA 953

QY 337 LeuGlyIleGlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyr 356
: : : : : : : : : : | | | | | : : : : : | | | | | : : : : :
Db 954 GTGGCACTGATGCTGATGTGGCTCAGCAATTTTCCGGAATCAATGGCATTTTCTACTAC 1013

QY 357 AlaSerSerIlePheLysAlaAlaGlyValThrAsnSerAspLeuAlaThrCysSerLeu 376
: : : : : : : : : : | | | | | : : : : : | | | | | : : : : :
Db 1014 TCAACCAAGCATTTTTCAGACGGCTGTATCAGCAAAACCTGTTTATGCAACCACTTGGAGTT 1073

QY 377 GlyAlaIleGlnValLeuAlaThrGlyValThrThrTrpLeuLeuAspArgAlaGlyArg 396
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 1074 GCGGCTGTAAACATGGTTTTCACCTGCTGCTGTATTCCTTGTGGAGAGGACGGCGGA 1133

QY 397 ArgIleLeuLeuIleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSerVal 416
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 1134 CGTTCTCTCTTCTAATTTGAATGAGTGGGATGTTTGTGTCATCTTTCATGTGAGTG 1193

QY 417 ValPhePheLeuLysAspAsnIleSerGlnAspSerAsnSerTyrTyrIleLeuThrMet 436
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 1194 GGACTTGTGCTG-----CTGAATAAGTTCTCTTGGATGAGTTAT 1232

QY 437 IleSerLeuValGlyIleValSerPheValIleThrPheSerPheGlyMetGlyAlaIle 456
: : : : : : : : : : | | | | | : : : : : | | | | | : : : : :
Db 1233 GTGAGCATGATAGCATCTTCTCTCTTGTGAGCTTCTTTGAAATTTGGGCCAGGCCGATC 1292

QY 457 ProTyrLeuMetMetSerGluIleLeuProValSerIleLysSerLeuGlySerIle 476
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 1293 CCCTGGTTTCATGGTGGCTGAGTTTTCAGTCAAGGACCACTGCTGCTGCTTGTAGCAATA 1352

QY 477 AlaThrLeuAlaAsnTyrLeuThrSerPheAlaIleThrMetThrAsnLeuMetLeu 496
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 1353 GCTGCATTCAGCAATTTGACCTGCAATTTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1412

QY 497 ThrTyrSerValGlyGlyThrPheLeuSerTyrMet-----ValValSerAlaPhe 513
: : : : : : : : : : | | | | | : : : : : | | | | | : : : : :
Db 1413 GACTTCTGT-----GGACCTTATGTTGTTTTCCTCTTTCCTGCTGAGTGTCTCTGCTGCT 1466

QY 514 ThrIleValPheValValLeuTyrValProGluThrLysGly 527
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 1467 ACCTG---TTTACATTTTAAAGTTCCAGAAACCAAGGA 1505

RESULT 2

US-09-919-172-79
; Sequence 79, Application US/09919172
; Patent No. 6673545
; GENERAL INFORMATION:
; APPLICANT: Faris, Mary
; APPLICANT: Turner, Christopher M.
; TITLE OF INVENTION: PROSTATE CANCER MARKERS
; FILE REFERENCE: PA-0036 US
; CURRENT APPLICATION NUMBER: US/09/919,172
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/222,469
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PERL Program
; SEQ ID NO 79
; LENGTH: 5227
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673545 410785.1
; NAME/KEY: unsure
; LOCATION: 4928, 4934, 4939, 4944, 4973, 4992
; OTHER INFORMATION: a, t, c, g, or other
US-09-919-172-79

Alignment Scores:
Pred. No.: 1.6e-50 Length: 5227
Score: 560.50 Matches: 153
Percent Similarity: 49.03% Conservative: 99
Best Local Similarity: 29.77% Mismatches: 169
Query Match: 20.17% Indels: 94
DB: 4 Gaps: 15

US-10-051-909-36 (1-553) x US-09-919-172-79 (1-5227)

QY 86 ThrSerSerMetAlaValLeuArgGluSerHisValSerAlaPheLeuCys---ThrLeu 104
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 30 ACTAGTACAATGACA-----GAAGATAAGGTCACCTGGGACCCCTGGTTTTCACCTGTC 80

Db 317 ACGTCCGACTGGATCGCGCGCTTCCACCATGCTTTCGCGCGCTCATCTTCTTCGCG 376
QY 179 GlyTrpLeuAlaIleSerPheAlaLysAspAlaSerPheLeuTyrMetGlyArgLeuLeu 198
Db 377 GCGCGCTCATCATGGGCTTCTCCGTCAACTACGCCATGCTCATGTTCTCGGCGCTTCG 436
QY 199 GluGlyPheGlyValGlyIleIleSerTyrThrValProValTyrIleAlaGluIleSer 218
Db 437 GCCGGCATCGCGGTGGGTACGCTCTCATGATCGCGCCCGTGAACACGGCGAGGTGCC 496
QY 219 ProGlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIle 238
Db 497 CCCGGTCTGCGCGTGGGTCTCTACATCTCTCCGGAGGTGTTCATCAACTTCGGCATC 556
QY 239 PheLeuAlaTyrLeu-----LeuGlyMetPheIleProTrpArg 251
Db 557 CTCCTCGGATATGTCTCCAACTTCGCCCTCGCCCGCTCTCCCTCCGCTCGGCTGGCGC 616
QY 252 LeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPhePheIle 271
Db 617 ATTATGCTCGGCATAGCGCGGTGCGTCTGCTCGGCTCATGGTGTCTCGGCATG 676
QY 272 ProGluSerProArgTrpLeuAlaLysMetAsnLeuThrGluAspCysGluThrSerLeu 291
Db 677 CCCGAGTCTCCCGGTGGTCTGTATGAAGGGCGGTCTCGCGACGCCAAGTTGTGCTT 736
QY 292 GlnValLeuArgGlyPheGluThrAspIleThrThrGluValAsnAspIleLysArgAla 311
Db 737 GCCAAGACGTCGACACACGCCGGAAGAGCGCGCGAGCGCATCGCGGACATTAAGACTGCC 796
QY 312 Val-----Ala 313
Db 797 GCCGGCATCCCTCTGGGCTCTGACGGGACGTGGTCCCGTGGCCCAAAACAAAGGAAGC 856
QY 314 SerSerSerLysArgThr-----ThrIleSerPheGlnGlu 325
Db 857 AGCGAGGAGAGCGGCTTTGAAGGACCTCATCTGTACCCGACCATAGCCATCGGCCAC 916
QY 326 LeuAsnGlnLysLysTyrArgThrProLeuLeuGlyIleGlyLeuValLeuGln 345
Db 917 ATC-----CTCATCGCGGAATCGGCATCCACTTCTTCCAG 952
QY 346 AsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSerIlePheLysAlaAlaGly 365
Db 953 CAGTCTTCGGGATCGACGCGCTCGTGTCTACAGCCCGCTAGTTTCAAGAGCGCGCGC 1012
QY 366 ValThrAsnSerAspLeu-----AlaThrCysSerLeuGlyAlaIleGlnValLeu 382
Db 1013 ATCAGGGCGGACAGCCGTCTCCGGCGGACACCGTGGCGGTTCGGGCGCCACCAATACG 1072
QY 383 AlaThrGlyValThrThrTrpLeuLeuAspArgAlaGlyArgArgIleLeuLeuIleIle 402
Db 1073 TTCATCTGTGGCCACCTTCTCTCCGACCGCATCCGCGCGCGCGCGCTGGTGTGACC 1132
QY 403 SerThrSerGlyMetThrLeuCysLeuLeuAlaValSerValPhePheLeuLysAsp 422
Db 1133 AGCAGGGCGGCGATGCTCTCTCTAGTGGGCTTCGCGACGGGCTCACCGTCTATCAGC 1192
QY 423 AsnIleSerGlnAspSerAsnSerTyrTyrIleLeuThrMetIleSerLeuValGlyIle 442
Db 1193 CGCCACCCGGACGAGAAGATCACCTGGGCGCTTCGCGACGGGCTCACCGTCTATCAGC 1246
QY 443 ValSerPheValIleThrPheSerPheGlyMetGlyAlaIleProTrpLeuMetMetSer 462
Db 1247 ATGGCCTACGTGGCTCTTCTTCCATCGGCTCGGCGCCCATCAGTGGGTGTACAGCTCG 1306
QY 463 GluIleLeuProValSerIleLysSerLeuGlyGlySerIleAlaThrLeuAlaAsnTrp 482
Db 1307 GAGATCTTCCCGTGCACGTGCGCGGCTGGGCTGTCTCCCTGGGCGTGGCGGTCAACCGC 1366
QY 483 LeuThrSerPheAlaIleThrMetThr---ThrAsnLeuMetLeuThrTrpSerValGly 501

Db 1367 CTGACCAGCGCGTGATCTCCATGACCTTCATTTCGCTGTCCAGGCCATGACCATCGGC 1426
QY 502 GlyThrPheLeuSerTyrMetValValSerAlaPheThrIleValPheValValLeuTrp 521
Db 1427 GCGCGCTTCTTCTCTTCGCGCGCATCGCTCATTCGCATGGGTGTTCTTCTTCGCCTAC 1486
QY 522 ValProGluThrLysGly 527
Db 1487 CTGCCGGAGACCGCGGC 1504
RESULT 4
US-09-291-922-21
; Sequence 21, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 21
; LENGTH: 2017
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-291-922-21
Alignment Scores:
Pred. No.: 7.33e-51 Length: 2017
Score: 557.00 Matches: 148
Percent Similarity: 49.01% Conservative: 99
Best Local Similarity: 29.37% Mismatches: 211
Query Match: 20.04% Indels: 46
DB: 3 Gaps: 9
US-10-051-909-36 (1-553) x US-09-291-922-21 (1-2017)
QY 80 SerSerPheAlaProGlyThrSerSerMetAlaValLeuArgGluSerHisValSer--- 98
Db 75 TCCTCTCTGTGACCAACCGAGATGGTTCGCGCGGTCCGCGGCGCGTCCGCGAGCGCGCGAAG 134
QY 99 -----AlaPheLeuCysThrLeuIleValAlaLeuGlyProfile 111
Db 135 AAGAAGGGCAACGTCCGGTTTCGCCCTTCGCTGCGCATCTCTCGCTCCATGACCTCCATC 194
QY 112 GlnPheGlyPheThrSerGlyPheSerSerProThrGlnAspAlaMetValArgAspLeu 131
Db 195 CTCCTCGGTACGATATCGGGGTGTGATGAGCGGGCGTCTCGTGTACATCAAGAGGACTTC 254
QY 132 AsnLeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyGlyMetVal 151
Db 255 AACATCAGTGACGGGAAGGTGGAGGTTCATGGGCATCTACTGAACCTCTACTCGCTCATC 314
QY 152 GlyAlaIleAlaSerGlyGlnMetAlaGluTyrIleGlyArgLysGlySerLeuMetIle 171
Db 315 GGCTCTTCGCGCGGGCGGACGTCGACTGGATCGGCGCGGTACACCATCGTGTTC 374
QY 172 AlaAlaIleProAsnIleIleGlyTrpLeuAlaIleSerPheAlaLysAspAlaSerPhe 191
Db 375 GCCCGCGTCATATTCTTCGCGGGGSGTTCCTCATGGGGTTCGCGGTCAACTACGCCATG 434
QY 192 LeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTyrThrValPro 211
Db 435 CTCATGTTCCGCGCGCTTCGTGGCGCGCATCGCGGTGGGTACGCGCTCATGATCGCGCGC 494
QY 212 ValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGln 231

Db 495 GTGTACACCGCGAGGTGTCGCCGGCGTTCGGCGCGTGGCTTCTGTACGTCGTTCCTCCGGAG 554

Qy 232 LeuSerValThrPheGlyIlePheLeuAlaTyrLeu-----Leu 244

Db 555 GTGTTTCATCAACTTCGGCATCCTGCTCGGTACGTCGAACTATGCTTTCTCCCGCTTG 614

Qy 245 GlyMetPheIleProTrpArgLeuLeuAlaValIleGlyAlaLeuProCysThrMetLeu 264

Db 615 CCGCTGAACCTCGGTGGCGCATCATGCTCGGCATCGCGCGCGCTCCGTGCTGCTC 674

Qy 265 IleProGlyLeuPhePheIleProGluSerProArgTrpLeuAlaLysMetAsnLeuThr 284

Db 675 GCGCTCATGTCGTCGGCATCGCGAGTCGCCGCGTGGTGGTATGATGAGGACGCCTC 734

Qy 285 GluAspCysGluThrSerLeuGlnValLeuArgGlyPheGluThrAspIleThrGlu 304

Db 735 GCGGACGCCAAGTGTGCTGGAGAAGACCTCCGACACGGCGGAGGAGCGCGGAGCGC 794

Qy 305 ValAsnAspIleLysArgAla-----ValAlaSer 314

Db 795 CTGGCCGACATCAAGCGCGCGGCATCCCTGTAGGAGCTCGACGGCGACGTGGTGACC 854

Qy 315 SerSerLysArgThrThrIleSerPheGlnGluLeuAsnGlnLysLysTyrArgThrPro 334

Db 855 GTCCCCAAGAGAGGAGCGGAACAGAGACGGGTGTGAAGAGCTCATCTGTCTCCCG 914

Qy 335 -----LeuLeuLeuGlyIleGlyLeuLeuValLeuGlnAsnLeu 347

Db 915 ACCCGGCCATCGCGCGCATCCTGTCTCGGGATCGGCATCCACTTCTTCCAGCATGC 974

Qy 348 SerGlyIleAsnGlyValLeuPheTyrAlaSerSerIlePheLysAlaAlaGlyValThr 367

Db 975 TTGGGCATTCACTCCGTCTCTTCTACAGCCCTCTCGTGTCAAGAGCCCCGGATTACG 1034

Qy 368 AsnSerAspLeu-----AlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThr 384

Db 1035 AAGCACAACACTTCTTGGGCACCACTTGGCCGTCGGTGTCAACCAAGAGCTTTTCATC 1094

Qy 385 GlyValThrThrTrpLeuLeuAspArgAlaGlyArgArgIleLeuLeuIleSerThr 404

Db 1095 TTGTTGGCGACTTCTTTCATCGACGGCGTCGGCGCGCGCGCTGTTGTCGGCAGCACG 1154

Qy 405 SerGlyMetThrLeuCysLeuLeuAlaValSerValValPhePheLeuLysAspAsnIle 424

Db 1155 GCGGGGATAATCCTCTCCCTCATCGGCCTCGCGCGCGGCTCACCGTC-----GTC 1205

Qy 425 SerGlnAspSerAsnSerTyrTyrIleLeuThrMet---IleSerLeuValGlyIleVal 443

Db 1206 GGCCAGACACCCGACGCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTC 1265

Qy 444 SerPheValIleThrPheSerPheGlyMetGlyAlaIleProTrpLeuMetMetSerGlu 463

Db 1266 GCCTACGTGCGCTTCTTCTCATCGGCCTTGGCCCCATCAGTGGGTGTACAGCTCGGAG 1325

Qy 464 IleLeuProValSerIleLysSerLeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeu 483

Db 1326 ATCTTCCCGCTCCAGGTGCGCGCGCTGGGCTGCTCGCTCGGCTGCGCGCAACCGCGTC 1385

Qy 484 ThrSerPheAlaIleThrMetThr---ThrAsnLeuMetLeuThrTrpSerValGlyGly 502

Db 1386 ACCAGCGCGTCATCTCCATGACCTTCTCTGTCGTGTCCAAGGCCATCACCATCGGCGC 1445

Qy 503 ThrPheLeuSerTyrMetValValSerAlaPheThrIleValPheValValLeuTrpVal 522

Db 1446 ACCTTCTTCTCTACTCCGGCATCGCGCGCTCGGCTGCTGCTGCTGCTTCTTACACCTACCTC 1505

Qy 523 ProGluThrLysGly**AsnSerArg-----GlyAspThrIlePhe 536

Db 1506 CCGGAGACCCGCGCGGACGCTGGAGGAGATGAGCAAGCTGTTGCGCGACACGCGCGCC 1565

Qy 537 ValSerLeuSer 540

Db 1566 GCCTCGGAATCA 1577

RESULT 5

US-09-023-655-1104
; Sequence 1104, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HERewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1104:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3915 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g183684
; US-09-023-655-1104

Alignment Scores:
Pred. No.: 9.46e-50 Length: 3915
Score: 551.50 Matches: 145
Percent Similarity: 48.95% Conservative: 87
Best Local Similarity: 30.59% Mismatches: 180
Query Match: 19.85% Indels: 62
DB: 4 Gaps: 13

US-10-051-909-36 (1-553) x US-09-023-655-1104 (1-3915)

QY 94 GluSerHisValSerAlaPheLeuCys-ThrLeuIleValAlaLeuGlyProIleGlnPh 113
Db 254 GAAGGTACACCCAGCTCTGTATATTGCCATCACAGTTGTACAAATCGGCTCTTTCCAAATT 313
QY 113 eGlyPheThrSerGlyPhe----- 119
Db 314 TGGCTACAACACTGGGGTTCATCAATGCTCTGAGAGATCATAAAGGAATTTATCAATAA 373
QY 120 -----SerSerProThrGlnAspAlaMetValArgAspLeu---As 132
Db 374 AACTTTGACGGACAAGGGAATGCCCCACCCCTCTGAGGTGCTGTCTGTCTCTGTGTC 433
QY 132 nLeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyGlyMetValGl 152

Db 659 TTTATTAATGGAGGATATTAAATTGGATACATATCAAACTATGTCATTTTCGAAGCTGACA 718
QY 246 MetPheIleProTyrArgLeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIle 265
Db 719 CTAAGCTGGGATGGCAATGATGCTTGGAGTTGGTGCATACCTTCGGTACTCCTAACA 778
QY 266 ProGlyLeuPhePheIleProGluSerProArgTyrPheLeuAlaLys----- 280
Db 779 GTAGGAGTGTGGCGATGCCGAGTCCCCAAGGTGGCTTGTGATGAGGGGTCTGTTGGGA 838
QY 281 -----MetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuArg 295
Db 839 GAGGCAAGAAAGTCTTAACAAAACCTCAGACAGCAAGGAAGGCCCAACTA----- 892
QY 296 GlyPheGluThrAspIleThrThrGluValAsnAspIleLysArgAlaVal-AlaSerSe 315
Db 893 -----AGGCTAGCGGAAATCAAAACAGCCGACGGATCCCC 928
QY 315 rSerLysArgThrThrIleSerPheGlnGluLeuAsnGln----- 328
Db 929 GAGAGTTGCAACGACGACGTCGTTTCAGGTAAATAAAACAAAGCAACGGTGAAGGTGTATGG 988
QY 329 -----LysLysTyrArgThrPro-----LeuLeuLeuGlyIleGl 340
Db 989 AAAGAGCTCTTCCTC-TATCCAACGCCCGCAATTCGTTCACATCGTAATCGCTGCCCTTGG 1047
QY 340 YLeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSerIl 360
Db 1048 TATTCACTTCTTCCAAACAAGCGTCGGCGGTAGACGCCGTCGTTTGTAGACGCCCAGGAT 1107
QY 360 ePheLysAlaAlaGlyValThrAsnSerAsp-----LeuAlaThrCysSerLeuGl 377
Db 1108 CTTCGAAAGGCTGGGATTACAAACGACACGACGATAAGCTTCTTGCACACCGTGGCGTTGG 1167
QY 377 YAlaIleGlnValLeuAlaThrGlyValThrThrTrpLeuLeuAspArgAlaGlyArgAr 397
Db 1168 ATTCTTAAGACCGGTTCATCTTGGCGGCTACGTTAGCTTGGACCGCGTGGTTCGTCG 1227
QY 397 gIleLeuLeuIleIleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSerValVa 417
Db 1228 TCCGTTGTTATTGTCFAGTGTCCGGCGCATGGTCTCTCGCTTCTCACGTTGCGATCAG 1287
QY 417 lPhePheLeuLysAspAsnIleSerGlnAspSerAsnSerTyrTyrIleLeuThrMetIl 437
Db 1288 CCTCACTGTTATTGATCAT---TCGGAGAGGAATAATGTTGGCGCGTTGGATCGAGCAT 1344
QY 437 eserLeuValGlyIleValserPheValIleThrPheSerPheGlyMetGlyAlaIlePr 457
Db 1345 AGCCATGGTG-----TTGGCTTACGTGGCCACGTTCTCCATCGGTGGCGGTCCCATCAC 1398
QY 457 oTrpLeuMetMetSerGluIleLeuProValSerIleLysSerLeuGlyGlySerIleAl 477
Db 1399 GTGGGTCTATAGTTCTGAGATCTTCCCGTTGAGGCTGGCGGCGCARGGTGGCGCGCGGG 1458
QY 477 aThrLeuAlaAsnTrpLeuThrSerPheAlaIleThrMetThr---ThrAsnLeuMetLe 496
Db 1459 AGTTGCGGTGAATAGGACCACTAGCGCGGTTGTCTCAATGACTTTTCTGTCCTCACTAG 1518
QY 496 uThrTrpSerValGlyThrPheLeuSerTyrMetValValSerAlaPheThrIleVa 516
Db 1519 AGCCATCACTATTGGTGGAGCTTCTTCCCTTATTGTGGCATTCGTACTGTGGGTGGAT 1578
QY 516 lPheValValLeuTrpValProGluThrLysGly 527
Db 1579 ATTCTTTTACACCGTCTTGCCTGAGACCCCGGGA 1612

RESULT 7

US-09-591-025-8
; Sequence 8, Application US/09591025
; Patent No. 6303373
; GENERAL INFORMATION:
; APPLICANT: Bogan, Jonathan S.
; APPLICANT: Lodish, Harvey F.

; TITLE OF INVENTION: Method of Measuring Plasma Membrane
; TITLE OF INVENTION: Targeting of GLUT4
; FILE REFERENCE: 0399.1210-004
; CURRENT APPLICATION NUMBER: US/09/591,025
; CURRENT FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/154,078
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: 60/138,237
; PRIOR FILING DATE: 1999-06-09
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 2592
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: modified GLUT4 containing myc tag sequences
US-09-591-025-8

Alignment Scores:

Pred. No.: 1.79e-47 Length: 2592
Score: 528.00 Matches: 131
Percent Similarity: 52.58% Conservative: 83
Best Local Similarity: 32.19% Mismatches: 165
Query Match: 19.00% Indels: 28
DB: 3 Gaps: 7

US-10-051-909-36 (1-553) x US-09-591-025-8 (1-2592)

QY 133 LeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyMetValGly 152
Db 562 CTCTCCGTGGCCATCTTTTC-----GTGGCGGCATGATTTC 600
QY 153 AlaIleAlaSerGlyGlnMetAlaGluTyrIleGlyArgLysGlySerLeuMetIleAla 172
Db 601 TCCTTCCTCATTTGGTATCATCTCTCAGTGGCTTGGAGAAAGGGCCATGCTGGTCAAC 660
QY 173 AlaIleProAsnIleIleGlyTrpLeuAlaIleSerPheAlaLysAspAlaSer----- 190
Db 661 AATGTCCTGGCGGTGCTGGGGGCGAGCCTCATGCGGCTGGCCACGCTGCTGCCTCCTAT 720
QY 191 ---PheLeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTyrThr 209
Db 721 GAAATGCTCATCTTGGACCATTCCTCATTTGGCGCCTACTCAGGGCTGACATCAGGSGTG 780
QY 210 ValProValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerVal 229
Db 781 GTGCCCATGTACGTGGGGGAGATTGCTCCCACTCACCTCGGGGGCGCCCTGGGGACGCTC 840
QY 230 AsnGlnLeuSerValThrPheGlyIlePheLeuAlaTyr-----Leu 243
Db 841 AACCAACTGGCCATTGTTATCGGCATTCTGATCGCCCGAGGTGCTGGGCTTGGAGTCCCTC 900
QY 244 LeuGlyMetPheIleProTyrArgLeuLeuAlaValIleGlyAlaLeuProCysThrMet 263
Db 901 CTGGGCACTGCCAGCCTGTGGCCACTGCTCCTGGGCTCAGAGTGTACCTGCTGCTGCTG 960
QY 264 LeuIleProGlyLeuPhePheIleProGluSerProArgTrpLeu---AlaLysMetAsn 282
Db 961 CAGCTGCTCCTGCTGCCCTTCTGTCCGAGAGCCCCCGTACCTCTACATCATCCAGAT 1020
QY 283 LeuThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPheGlu-----ThrAsp 300
Db 1021 CTCGAGGGGCTGCCAGAAAGAGTGTGAAGCGCTGACAGGCTGGGCGGATGTTTCTGGA 1080
QY 301 IleThrThrGluValAsnAspIleLysArgAlaValAlaSerSerLysArgThrThr 320
Db 1081 GTGCTGGCTGAGCTGAAGGATGAGAAGCGGAAGCTG-----GAGCGTGACGGCCACTG 1134
QY 321 IleSerPheGlnGluLeuAsnGlnLysLysTyrArgThrProLeuLeuGlyIleGly 340
Db 1135 TCCCTGCTCCAGCTCCTGGGCGAGCCGTCACCCACCGGCGAGCCCTGATCATTCGGGTCTG 1194


```
Db 1160 ACCGTCGGGGGCAATCACCGGACACCAAGGTCGGGTGGCGGTC-----GCCCTGTGC 1213
QY 439 LeuValGlyIleValSerPheValIleThrPheSerPheGlyMetGlyAlaIleProTrp 458
Db 1214 ATCGCGTCAACCCCTGTCTACATCGCCTTCTTCTCATCGGCTCGGGCCCATCACGGGC 1273
QY 459 LeuMetMetSerGluIleLeuProValSerIleLysSerLeuGlyGlySerIleAlaThr 478
Db 1274 GTGTACACCTCGGAATATTCCCGCTGCAGGTGCGCGCGCTGGGCTTCGGCGTGGGTGTG 1333
QY 479 LeuAlaAsnTrpLeuThrSerPheAlaIleThrMetThr---ThrAsnLeuMetLeuThr 497
Db 1334 CCGAGCAACCGCGTCACAGCGCGCTCATCTCCATGACCTTCCTGTCCCTTCCAGGCC 1393
QY 498 TrpSerValGlyThrPheLeuSerTyrMetValValSerAlaPheThrIleValPhe 517
Db 1394 ATCACCATCGGGCGGAGCTTCTTCTACTCCGGCATCGCCGGTCTGGGTTTTC 1453
QY 518 ValValLeuTrpValProGluThrLysGly 527
Db 1454 TTCTTACGTGCTCCCGGAGACACGGGC 1483

RESULT 9
US-09-894-927B-8
; Sequence 8, Application US/09894927B
; Patent No. 6632924
; GENERAL INFORMATION:
; APPLICANT: Bogan, Jonathan S.
; APPLICANT: Lodish, Harvey F.
; TITLE OF INVENTION: Method of Measuring Plasma Membrane
; TITLE OF INVENTION: Targeting of GLUT4
; FILE REFERENCE: 0399.1210-005
; CURRENT APPLICATION NUMBER: US/09/894,927B
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US 09/591,025
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: US 60/154,078
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: US 60/138,237
; PRIOR FILING DATE: 1999-06-09
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 2592
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: modified GLUT4 containing myc tag sequences
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2592)
US-09-894-927B-8

Alignment Scores:
Pred. No.: 4.94e-47 Length: 2592
Score: 524.00 Matches: 131
Percent Similarity: 52.33% Conservative: 82
Best Local Similarity: 32.19% Mismatches: 166
Query Match: 18.86% Indels: 28
DB: 4 Gaps: 7

US-10-051-909-36 (1-553) x US-09-894-927B-8 (1-2592)
QY 133 LeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyGlyMetValGly 152
Db 562 CTCTCCGTGGCCATCTTTTCC-----GTGGCGGCGCATGATTCC 600
QY 153 AlaIleAlaSerGlyGlnMetAlaGluTyrIleGlyArgLysGlySerLeuMetIleAla 172
Db 601 TCCTTCTCATTTGGTATCATCTCTCAGTGGCTTGGAGGAAAGGGCCATGCTGGTCAAC 660
QY 173 AlaIleProAsnIleIleGlyTrpLeuAlaIleSerPheAlaLysAspAlaSer----- 190
```

```
Db 661 AATGCTCTGGCGGTGCTGGGGGGAGCCTCATGGGCTTGCCAAACGCTGCTGCCTCTAT 720
QY 191 ---PheLeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTyrThr 209
Db 721 GAAATGCTCATCTTGGACGATTCCTCATTTGGCGCTACTCAGGGCTGACATCAGGGCTG 780
QY 210 ValProValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerVal 229
Db 781 GTGCCCATGTACGTGGGGGAGATTGCTCCCACTCACCTGGGGGGCGCCCTGGGGACGCTC 840
QY 230 AsnGlnLeuSerValThrPheGlyIlePheLeuAlaTyr-----Leu 243
Db 841 AACCAACTGGCCATTGTTATCGGCATTCTGATCGCCAGGTGCTGGGCTTGGAGTCCCTC 900
QY 244 LeuGlyMetPheIleProTrpArgLeuLeuAlaValIleGlyAlaLeuProCysThrMet 263
Db 901 CTGGGCACTGCCAGCCCTGTGGCCACTGCTCCTGGCCCTCACAGTGTACCTGCCCTCCTG 960
QY 264 LeuIleProGlyLeuPhePheIleProGluSerProArgTrpLeu---AlaLysMetAsn 282
Db 961 CAGCTGGTCTGCTGCCCTTCTGTCCGAGAGCCCCCGCTACCTCTACATCATCCAGAAT 1020
QY 283 LeuThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPheGlu-----ThrAsp 300
Db 1021 CTCGAGGGGCTGCCAGAAAGAGCTCTGAAGCGCTGACAGGCTGGCCGATGTTTCTGGA 1080
QY 301 IleThrThrGluValAsnAspIleLysArgAlaValAlaSerSerLysArgThrThr 320
Db 1081 GTGCTGGCTGAGCTGAAGGATGAGAAAGCGGAAGCTG-----GAGCGTGAGCGGCCACTG 1134
QY 321 IleSerPheGlnGluLeuAsnGlnLysLysTyrArgThrProLeuLeuGlyIleGly 340
Db 1135 TCCTGCTCCAGCTCCTGGGCAAGCGGTACCCACCGGCGAGCCCTGATCATTTGGGTCGTG 1194
QY 341 LeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSerIle 360
Db 1195 CTGCACTGAGCCAGCAGCTCTCTGGCATCAATGCTGTTTCTATTATTTCGACCGCATC 1254
QY 361 PheLysAlaAlaGlyValThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIleGln 380
Db 1255 TTCGAGACAGCAGGGGTAGGCCAGCCTGCCTATGCCACCACATAGGAGCTGGTGTGTCAC 1314
QY 381 ValLeuAlaThrGlyValThrTrpLeuLeuAspAlaGlyArgAlaGlyLeuLeu 400
Db 1315 ACAGTCTTCACCTTGGTCTCGGTGTTGTTGGTGGAGCGGGCGGCGCGCATCCAT 1374
QY 401 IleIleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSerValPhePheLeu 420
Db 1375 CTCCTGGGCTTGGCGGCGCATGTGGCTGTGCCATCCTGATGACTGTGGCTCTGCTCCTG 1434
QY 421 LysAspAsnIleSerGlnAspSerAsnSerTyrTyrIleLeuThrMetIleSerLeuVal 440
Db 1435 CTGGAGCGAGTTCCA-----GCCATGAGCTACGTCTCCATTGTG 1473
QY 441 GlyIleValSerPheValIleThrPheSerPheGlyMetGlyAlaIleProTrpLeuMet 460
Db 1474 GCCATCTTTGGCTTGGTGGCATTTTGGAGATTGGCCCTGGCCCCCATTCCTTGGTTTCATC 1533
QY 461 MetSerGluIleLeuProValSerIleLysSerLeuGlyGlySerIleAlaThrLeuAla 480
Db 1534 GTGGCCGAGCTCTTCAGCCAGGACCCCGCCGCGCAGCCATGGCTGGGCTGGTTTCTCC 1593
QY 481 AsnTrpLeuThrSerPheAlaIleThrMetThrThrAsnLeuMetLeuThrTrpSerVal 500
Db 1594 AACTGGACGAGCAACTTCATCATTTGGCATGGGTTTCCAGTATGTTGGGAGGCTATGGGG 1653
QY 501 GlyGlyThrPheLeuSerTyrMetValValSerAlaPheThrIleValPheValLeu 520
Db 1654 CCCTACGCTTCTCTTCTATTGCGTCTCTCTGCTGGGCTCTCTTCTTCTTCTTCTTCTT 1713
QY 521 TrpValProGluThrLysGly 527
```


Db 1714 AGAGTACCTGAAACTCGAGGC 1734

RESULT 10

US-09-643-597-135

; Sequence 135, Application US/09643597

; Patent No. 6426072

; GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong

; APPLICANT: Fan, Liqun

; APPLICANT: Kalos, Michael D.

; APPLICANT: Bangur, Chaitanya S.

; APPLICANT: Hosken, Nancy

; APPLICANT: Fanger, Gary R.

; APPLICANT: Li, Samuel X.

; APPLICANT: Wang, Aijun

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Henderson, Robert A.

; APPLICANT: McNeill, Patricia D.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; FILE REFERENCE: 210121.455C11

; CURRENT APPLICATION NUMBER: US/09/643,597

; CURRENT FILING DATE: 2000-08-21

; NUMBER OF SEQ ID NOS: 369

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 135

; LENGTH: 2856

; TYPE: DNA

; ORGANISM: Homo sapien

US-09-643-597-135

Alignment Scores:

Pred. No.:	6.64e-47	Length:	2856
Score:	523.50	Matches:	151
Percent Similarity:	45.91%	Conservative:	96
Best Local Similarity:	28.07%	Mismatches:	229
Query Match:	18.84%	Indels:	62
DB:	4	Gaps:	12

US-10-051-909-36 (1-553) x US-09-643-597-135 (1-2856)

QY 27 ProGlyArgProAlaSerGluLeuArgThrArgValMetGlyGlySer-AsnArgGl 46

Db 14 CCGAGTGAGCACGCCAGGAGCAGGAGACCAACACGCGGGGTGGAGTCAGAGTCGAG 73

QY 46 yGlyAla-----GlyAlaGlyGluGluSerGlySerAspHisAspGlyValLeuAr 63

Db 74 TGGGAGTCCCGGACCGGAGCAGCAGCCTGAGCGGAGAGCGCGCTCGCACGCCGCTGC 133

QY 63 gArgProLeuLeuAsnThrGlySerTrpTyrArgMetSerSerArgGlnSerSerPheAl 83

Db 134 CCACCCGCGTACCCGG-----CGCAGCCAGAGCCACCGAGCGCGCTGC 178

QY 83 aProGlyThrSerSerMetAlaValLeuArgGluSerHisValSerAlaPheLeuCysTh 103

Db 179 CATGGAGCCCGACGACGAAGAAGCTGACGGGTGCG-----CTCATGCTGGCTGT 226

QY 103 rLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGlyPheSerSerProth 123

Db 227 GGGAGGAGCAGTGTGGCTCCCTGCAGTTGGCTACAACTGAGTCATCAATGCCCC 286

QY 123 rGlnAspAlaMetValArgAspLeuAsn----- 132

Db 287 CCAGAAGGTGATCGAGGAGTTCTACACGACATGGGTCCACCGCTATGGGAGAGCAT 346

QY 133 -----LeuSerIleSerGluPheSerAlaPheGl 142

Db 347 CCTGCCACCACGCTCACCACGCTCTGGTCCCTCTCAGTGGCCATCTTTTCT----- 398

QY 142 ySerLeuSerAsnValGlyMetValGlyAlaIleAlaSerGlyGlnMetAlaGluTy 162

Db 399 -----GTTGGGGGCATGATTGGCTCCTTCTCTGTGGGCTTTTTCGTTAACCG 445

QY 162 rIleGlyArgLysGlySerLeuMetIleAlaAlaIleProAsnIleIleGlyTrpLeuAl 182

Db 446 CTTTGGCCGCGGAATTCAATGCTGATGATGAACCTGCTGGCCTCGTGTCCGCCGTGCT 505

QY 182 aIleSerPheAla-----LysAspAlaSerPheLeuTyrMetGlyArgLeuLeuGl 199

Db 506 CATGGGCTTCTCGAAACTGGGCAAGTCTTTGAGATGCTGATCGTGGGCGCGCTTCATCAT 565

QY 199 uGlyPheGlyValGlyIleIleSerTyrThrValProValTyrIleAlaGluIleSerPr 219

Db 566 CGGTGTGTACTGCGGCTTGACACACAGGCTTCGTGCCCATGTATGTGGGTGAAGTGTCAAC 625

QY 219 oGlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIlePh 239

Db 626 CACAGCCTTTCTGGGCGCTGGGCAACCTGCACAGCTGGGCATCGTCTGGCATCCT 685

QY 239 eLeuAlaTyrLeuLeuGlyMetPheIle-----ProTyrArgLeuLe 253

Db 686 CATGCCCGAGGTGTTGGCCTGGACTCCATCATGGGCAACAAGGACCTGTGGCCCTGCT 745

QY 253 uAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPhePheIleProGl 273

Db 746 GCTGAGCATCATCTTCATCCCGGCTTCGTGAGTGCATCGTGTGCCCTTCTGCCCGCA 805

QY 273 uSerProArgTrpLeuAlaLysMetAsnLeuThrGluAsp-----CysGluThrSerLe 291

Db 806 GAGTCCCGCTTCTGCTC---ATCAACCGCAACGAGGAGAACCGGGCCAAAGAGTGTGCT 862

QY 291 uGlnValLeuArgGlyPheGluThrAspIleThrThrGluValAsnAspIleLysArgAl 311

Db 863 AAAGAAGCTGCGCGG---ACAGCTGACGTGACCCATGACCTGAGGAGATGAAGGAAGA 919

QY 311 aValAlaSerSerLysArgThrThrIleSerPheGlnGluLeu---AsnGlnLysLy 330

Db 920 GAGTCGGCAGATCATCGGAGAGAGAGTCAACATCCTGGAGTGTTCGCTCCCGCGC 979

QY 330 sTyrArgThrProLeuLeuGlyIleGlyLeuValLeuValLeuGlnAsnLeuSerGlyIl 350

Db 980 CTACCGCCAGCCCATCTCATCGTGTGGTGTGACGTGTCCAGCAGCTGTCTGGCAT 1039

QY 350 eAsnGlyValLeuPheTyrAlaSerSerIlePheLysAlaAlaGlyValThrAsnSerAs 370

Db 1040 CAACGCTGTCTTCTATTACTCCAGCAGCATCTTCGAGAAGCGGGGTGCAGCAGCCTGT 1099

QY 370 pLeuAlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThrGlyValThrTrpLe 390

Db 1100 GTATGCCACCATTTGGTCCGGTATCGTCAACACAGGCTTCACCTGCTGTGCTGTGT 1159

QY 390 uLeuAspArgAlaGlyArgIleLeuLeuIleIleSerThrSerGlyMetThrLeuCy 410

Db 1160 GGTGGAGCGAGCAGCGCGGACCCCTGACCTCATAGGCTCGTGGCATGGCGGTTG 1219

QY 410 sLeuLeuAlaValSerValPhePheLeuLysAspAsnIleSerGlnAspSerAsnSe 430

Db 1220 TGCCATACTCATGACCATCGCGCTAGCAGTGTGGAGCAGCTACCCCTGG----- 1268

QY 430 rTyrTyrIleLeuThrMetIleSerLeuValGlyIleValSerPheValIleThrPheSe 450

Db 1269 -----ATGTCCTATCTGAGCATCGTGGCCATCTTTGGCTTTGTGGCCTTCTTTGA 1318

QY 450 rPheGlyMetGlyAlaIleProTrpLeuMetMetSerGluIleLeuProValSerIleLy 470

Db 1319 AGTGGGTCTCTGGCCCCATCCCATGGTTTCATCGTGGCTGAACCTTCAGCCAGGGTCCACG 1378

QY 470 sSerLeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeuThrSerPheAlaIleThrMe 490

Db 1379 TCCAGCTGCCATTCGCGTTCAGGCTTCCAACTGAGCTTCAGCCAGGGTCCACG 1438

QY 490 tThrThrAsnLeuMetLeuThrTrpSerValGlyGlyThrPheLeuSerTyrMetValva 510

Db 1439 GTGCTTCCAGTATGTGGAGCAACTGTGTGGTCCCTACGTCTTCATCATCTTCACGTGTCT 1498

QY 510 lSerAlaPheThrIleValPheValLeuTrpValProGluThrLysGly 527

```

Db      1499 CCTGGTTCTGTTTTCATCTTCACTACTTCAAAGTTCTCGAGACTAAAGGC 1550
          :   ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
RESULT 11
US-09-480-884A-135
; Sequence 135, Application US/09480884A
; Patent No. 6482597
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C6
; CURRENT APPLICATION NUMBER: US/09/480,884A
; CURRENT FILING DATE: 2001-08-27
; NUMBER OF SEQ ID NOS: 330
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 135
; LENGTH: 2856
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-480-884A-135

Alignment Scores:
Pred. No.:        6.64e-47           Length:       2856
Score:            523.50             Matches:     151
Percent Similarity: 45.91%           Conservative: 96
Best Local Similarity: 28.07%         Mismatches:  229
Query Match:      18.84%              Indels:      62
DB:                4                  Gaps:        12

US-10-051-909-36 (1-553) x US-09-480-884A-135 (1-2856)
QY      27 ProGlyArgProAlaSerGluLeuArgThrArgValMetGlyGlyGlySer-AsnArgGl 46
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      14 CCGAGTGACACGCCAGGAGCAGGAGCACGAGCCGTGACGGGAGAGCGCCGCTCGCAGTCGCGAGT 73
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      46 yGlyAla-----GlyAlaGlyGluGluSerGlySerAspHisAspGlyValLeuAr 63
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      74 TGGGAGTCCCCGACCAGGAGCACGAGCCGTGACGGGAGAGCGCCGCTCGCAGTCGCGAGT 133
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      63 gArgProLeuLeuAsnThrGlySerTrpTyArgMetSerSerArgGlnSerSerPheAl 83
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      134 CCACCCGCGTACCCGG-----CGCAGCCAGAGCCACCAGCGCAGCGCGCTGC 178
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      83 aproGlyThrSerSerMetAlaValLeuArgGluSerHisValSerAlaPheLeuCysTh 103
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      179 CATGGAGCCCAGCAGCAAGAAGCTGACGGGTCCG-----CTCATGCTGGCTGT 226
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      103 rLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGlyPheSerSerProTh 123
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      227 GGGAGGAGCAGTGCTTGGCTCCCTGCAGTTTGGCTACAACACTGGAGTCATCAATGCCCC 286
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      123 rGlnAspAlaMetValArgAspLeuAsn-----LeuSerIleSerGluPheSerAlaPheGl 132
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      287 CCAGAGGTGATCGAGGAGTTCTACAACCAGACATGGTCCACCGCTATGGGAGAGCAT 346
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      133 -----LeuSerIleSerGluPheSerAlaPheGl 142
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      347 CCTGCCCCACACGCTCACCACGCTCTGGTCCCTCTCAGTGGCCATCTTTTCT----- 398
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      142 ySerLeuSerAsnValGlyGlyMetValGlyAlaIleAlaSerGlyGlnMetAlaGluTy 162
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      399 -----GTTGGGGCATGATTGGTCTCTTCTGTGGGCCCTTTTCGTTAACCG 445
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      162 rIleGlyArgLysGlySerLeuMetIleAlaAlaIleProAsnIleIleGlyTrpLeuAl 182
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      446 CTTTGGCCGCGGAATTCAATGCTGATGTAACCTGCTGGCCTTCGTGTCCGCCGTGCT 505
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      182 aileSerPheAla-----LysAspAlaSerPheLeuTyrMetGlyArgLeuLeuGl 199

```

; Sequence 135, Application US/09542615A
; Patent No. 6518256
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; FILE REFERENCE: 210121.455C8
; CURRENT APPLICATION NUMBER: US/09/542,615A
; CURRENT FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 350
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 135
; LENGTH: 2856
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-542-615A-135

Alignment Scores:
pred. No.: 6.64e-47 Length: 2856
Score: 523.50 Matches: 151
Percent Similarity: 45.91% Conservative: 96
Best Local Similarity: 28.07% Mismatches: 229
Query Match: 18.84% Indels: 62
DB: 4 Gaps: 12

US-10-051-909-36 (1-553) x US-09-542-615A-135 (1-2856)

QY 27 ProGlyArgProAlaSerGluLeuArgThrArgValMetGlyGlyGlySer-AsnArgG1 46
DB 14 CCGAGTGAGCACGCCAGGAGCAGGAGACCAACACGACGGGGTGGAGTCAGAGTCGCAG 73
QY 46 YGlyAla-----GlyAlaGlyGluGluSerGlySerAspHisAspGlyValLeuAr 63
DB 74 TGGGAGTCCCGGACCGGAGCAGCAGCGCTGAGCGGGAGAGCGCGCTCGCACGCCCGTCG 133
QY 63 gArgProLeuLeuAsnThrGlySerTrpTyrArgMetSerSerArgGlnSerSerPheAl 83
DB 134 CCACCCGCGGTACCCGG-----CGCAGCCAGAGCACCAGCGCAGCGCGCTGC 178
QY 83 aProGlyThrSerSerMetAlaValLeuArgGluSerHisValSerAlaPheLeuCysTh 103
DB 179 CATGGAGCCCGCAGCAGCAAGAGCTGACGGGTGCG-----CTCATGCTGGCTGT 226
QY 103 rLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGlyPheSerSerProTh 123
DB 227 GGGAGGAGCAGTGCTGGCTCCCTGCGAGTTGGCTACAACTGGAGTCATCAATGCCCC 286
QY 123 rGlnAspAlaMetValArgAspLeuAsn-----132
DB 287 CCAGAAGGTGATCGAGGAGTTCTACACACAGACATGGGTCCACCGCTATGGGGAGAGCAT 346
QY 133 -----LeuSerIleSerGluPheSerAlaPheG1 142
DB 347 CCTGCCACACCGCTCACCAGCTCTGGTCCCTCTCAGTGGCCATCTTTCT----- 398
QY 142 ySerLeuSerAsnValGlyGlyMetValGlyAlaIleAlaSerGlyGlnMetAlaGluTy 162
DB 399 -----GTTGGGGGATGATGGCTCTCTCTCTGGGCCCTTTTCTGTTAACCG 445
QY 162 rIleGlyArgLysGlySerLeuMetIleAlaAlaIleProAsnIleIleGlyTrpLeuAl 182
DB 446 CTTTGGCCCGCGGAATTCATGCTGATGATGAACCTGCTGGCCCTCGTGTCCGCCGTGCT 505
QY 182 aIleSerPheAla-----LysAspAlaSerPheLeuTyrMetGlyArgLeuLeuG1 199
DB 506 CATGGGCTTCTCGAAACTGGGCAAGTCTTTTGAGATGCTGATCTGGGCCGCTTTCATCAT 565
QY 199 uGlyPheGlyValGlyIleIleSerTyrThrValProValTyrIleAlaGluIleSerPr 219

DB 566 CGGTGTGTACTCGGCGCTGACCACAGGCTTCGTGCCCCATGTATGTGGTGAAGTGTCAACC 625
QY 219 oGlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIlePh 239
DB 626 CACAGCCCTTTCGTGGGGCCCTGGGCAACCTGCACAGCTGGGATCGTCTGGCATCCT 685
QY 239 eLeuAlaTyrLeuLeuGlyMetPheIle-----ProTrpArgLeuLe 253
DB 686 CATGCCCGCAGGTGTTCCGCTGGACTCCATCATGGGCAACAGGACCTGTGGCCCTGCT 745
QY 253 uAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPhePheIleProG1 273
DB 746 GCTGAGCATCATCTTCATCCCGGCCCTGCTGAGTGCATCGTGTGCCCTTCTGCCCGCA 805
QY 273 uSerProArgTrpLeuAlaLysMetAsnLeuThrGluAsp-----CysGluThrSerLe 291
DB 806 GAGTCCCGCTTCCTGCTC--ATCAACCGCAACGAGGAGAACCGGGCCCAAGAGTGTGCT 862
QY 291 uGlnValLeuArgGlyPheGluThrAspIleThrGluValAsnAspIleLysArgAl 311
DB 863 AAAGAAGCTGCGCGGG--ACAGCTGACCTGACCCATGACCTGAGGAGATGAAGGAAGA 919
QY 311 aValAlaSerSerLysArgThrThrIleSerPheGlnGluLeu---AsnGlnLysLy 330
DB 920 GAGTCGGCAGATGATCGGGAGAGAGAGTCAACATCTCGAGCTGTTCGGTCCCCCGC 979
QY 330 sTyrArgThrProLeuLeuGlyIleGlyLeuValLeuValLeuGlnAsnLeuSerGlyI1 350
DB 980 CTACCCGCCAGCCCATCTCATCGCTGTGTGTGCTGAGCTGTCCAGCAGCTGTCTGGCAT 1039
QY 350 eAsnGlyValLeuPheTyrAlaSerSerIlePheLysAlaAlaGlyValThrAsnSerAs 370
DB 1040 CAACGCTGTCTTCTATTACTCCACGAGCATCTTCGAGAGGCGGGGTGCAGCAGCCTGT 1099
QY 370 pLeuAlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThrGlyValThrTrpLe 390
DB 1100 GTATGCCACCATGGCTCCGGTATCGTCAACACGGCTTCACCTGCTGCTGCTGTTGT 1159
QY 390 uLeuAspArgAlaGlyArgArgIleLeuLeuIleIleSerThrSerGlyMetThrLeuCy 410
DB 1160 GGTGGAGCGAGCAGCGCGGAGCCCTGCACCTCATAGGCCCTGCTGGCATGGCGGGTTG 1219
QY 410 sLeuLeuAlaValSerValValPhePheLeuLysAspAsnIleSerGlnAspSerAsnSe 430
DB 1220 TGCCATACCTCATGACCATCGCGCTAGCAGCTGTGGAGCAGCTACCCCTGG----- 1268
QY 430 rTyrTyrIleLeuThrMetIleSerLeuValGlyIleValSerPheValIleThrPheSe 450
DB 1269 -----ATGTCTCTATCTGAGCATCGTGGCCATCTTTGGCTTTGTGGCCTTCTTTGA 1318
QY 450 rPheGlyMetGlyAlaIleProTrpLeuMetMetSerGluIleLeuProValSerIleLy 470
DB 1319 AGTGGGTCTCTGCCCCCATCCCATGGTTCATCTGCTGGCTGAACCTTCAGCCAGGGTCCACG 1378
QY 470 sSerLeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeuThrSerPheAlaIleThrMe 490
DB 1379 TCCAGCTGCCATTCGCTGAGGCTTCTCCAACTGGACCTCAAAATTTTCATTGTGGGCAT 1438
QY 490 tThrThrAsnLeuMetLeuThrTrpSerValGlyGlyThrPheLeuSerTyrMetValVa 510
DB 1439 GTGCTTCCAGTATGGAGCAACTGTGTGGTCCCTACGTCTTCATCTTCATCTTACGTGCT 1498
QY 510 lSerAlaPheThrIleValPheValValLeuTrpValProGluThrLysGly 527
DB 1499 CCTGGTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1550

RESULT 13

US-09-606-421B-135
; Sequence 135, Application US/09606421B
; Patent No. 6531315
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong


```

; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C9
; CURRENT APPLICATION NUMBER: US/09/606,421B
; CURRENT FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 135
; LENGTH: 2856
; TYPE: DNA
; ORGANISM: Homo sapien
; US-09-606-421B-135

Alignment Scores:
Pred. No.: 6.64e-47 Length: 2856
Score: 523.50 Matches: 151
Percent Similarity: 45.91% Conservative: 96
Best Local Similarity: 28.07% Mismatches: 229
Query Match: 18.84% Indels: 62
DB: 4 Gaps: 12

US-10-051-909-36 (1-553) x US-09-606-421B-135 (1-2856)
QY 27 ProGlyArgProAlaSerGluLeuArgThrArgValMetGlyGlyGlySer-AsnArgG1 46
Db 14 CCGAGTGACACGCGCCAGGACGAGAGACCAACGACGGGGTCCGAGTCAGAGTCGCAG 73
QY 46 yGlyAla-----GlyAlaGlyGluGluSerGlySerAspHisAspGlyValLeuAr 63
Db 74 TGGAGTCCCGGACCGGACGAGCCTGAGCGGGAGAGCGCGCTCGCACGCCGCTCG 133
QY 63 gArgProLeuLeuAsnThrGlySerTrpTyrArgMetSerSerArgGlnSerSerPheAl 83
Db 134 CCACCCGCGTACCCGG-----CGCAGCCAGAGCCACCAGCGCAGCGCTGC 178
QY 83 aProGlyThrSerSerMetAlaValLeuArgGluSerHisValSerAlaPheLeuCysTh 103
Db 179 CATGAGCCCGACGACGACGAGAGCTGACGGGTGCG-----CTCATGCTGGCTGT 226
QY 103 rLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGlyPheSerProTh 123
Db 227 GGGAGGAGCAGTGTCTGGCTCCCTGCGAGTTTGGCTACACACTGGAGTCATCAATGCCCC 286
QY 123 rGlnAspAlaMetValArgAspLeuAsn----- 132
Db 287 CCAGAAGGTGATCGAGGAGTTCTACAAACAGACATGGGTCCACCGCTATGGGAGAGCAT 346
QY 133 -----LeuSerIleSerGluPheSerAlaPheG1 142
Db 347 CCTGCCACACCGCTCACACGCTGCTGCTCCCTCTCATGTGGCCATCTTTTCT----- 398
QY 142 ySerLeuSerAsnValGlyGlyMetValGlyAlaIleAlaSerGlyGlnMetAlaGluTy 162
Db 399 -----GTTGGGGCATGATTGGCTCCTTCTGTGGGCTTTTTCGTTAACCG 445
QY 162 rIleGlyArgLysGlySerLeuMetIleAlaAlaIleProAsnIleIleGlyTrpLeuAl 182
Db 446 CTTTGGCCGGCGGAATTCAATGTGATGATGAACCTGTGGCCTTCGTGTCCGCCGTGCT 505
QY 182 aIleSerPheAla-----LysAspAlaSerPheLeuTyrMetGlyArgLeuLeuG1 199
Db 506 CATGGCTTCTCGAAACTGGGCAAGTCCTTTGAGATGCTGATCCTGGGCCGCTTCATCAT 565
QY 199 uGlyPheGlyValGlyIleIleSerTyrThrValProValTyrIleAlaGluIleSerPr 219
```

```

Db 566 CGGTGTGTACTGCGGCGCTGACCACAGGCTTCGTGCCCATGTATGTGGTGAAGTGTACCC 625
QY 219 oGlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIlePh 239
Db 626 CACAGCCTTTTCGTGGGGCCCTGGGCACCCCTGCACACAGTGGGCATCGTCTCGGCATCCT 685
QY 239 eLeuAlaTyrLeuLeuGlyMetPheIle-----ProTyrArgLeuLe 253
Db 686 CATGCCCCAGGTGTTCGGCCCTGGACTCCATCATGGGGCAACAAGGACCTGTGGCCCTGCT 745
QY 253 uAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPhePheIleProG1 273
Db 746 GCTGAGCATCATCTTCATCCCGGCCCTGCTGCAGTGCATGCTGCTGCCCTTCTGCCCCGA 805
QY 273 uSerProArgTrpLeuAlaLysMetAsnLeuThrGluAsp-----CysGluThrSerLe 291
Db 806 GAGTCCCGCTTCCTGTCTC---ATCAACCGCAACGAGGAGAACCGGGCCAAAGAGTGTGCT 862
QY 291 uGlnValLeuArgGlyPheGluThrAspIleThrThrGluValAsnAspIleLysArgAl 311
Db 863 AAAGAAAGCTGCGCGG---ACAGCTGACGTGACCCATGACCTGCAGGAGATGAAGGAAGA 919
QY 311 aValAlaSerSerSerLysArgThrThrIleSerPheGlnGluLeu---AsnGlnLysLy 330
Db 920 GAGTCCGCGAGATGATCGGGGAGAGAAGGTACCATCTCTGGAGTGTTCGCTCCCGCTCCCGCG 979
QY 330 sTyrArgThrProLeuLeuGlyIleGlyLeuLeuValLeuGlnAsnLeuSerGlyI1 350
Db 980 CTACCGCCAGCCCATCCTCATCGCTGTGGTGTGCTGCTGCTGCCAGCAGCTGTCTGGCAT 1039
QY 350 eAsnGlyValLeuPheTyrAlaSerSerIlePheLysAlaAlaGlyValThrAsnSerAs 370
Db 1040 CAACGCTGTCTTCTATTACTCCAGCAGCATCTTCGAGAAGCGGGGGTGCAGCAGCCTGT 1099
QY 370 pLeuAlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThrGlyValThrTrpLe 390
Db 1100 GTATGCCACCATTTGGCTCCGGTATCGTCAACACGGCCTTCACTGCTGCTGCTGTTGT 1159
QY 390 uLeuAspArgAlaGlyArgArgIleLeuLeuIleIleSerThrSerGlyMetThrLeuCy 410
Db 1160 GGTGAGCGAGCAGGCGCGGGACCCTGCACCTCATAGGCTCGTGGCATGGCGGGTTG 1219
QY 410 sLeuLeuAlaValSerValPhePheLeuLysAspAsnIleSerGlnAspSerAsnSe 430
Db 1220 TGCCATACTCATGACCATCGCGCTAGCAGTGTGGAGCAGCTACCTGG----- 1268
QY 430 rTyrTyrIleLeuThrMetIleSerLeuValGlyIleValSerPheValIleThrPheSe 450
Db 1269 -----ATGTCTTATCTGAGCATCGTGGCCATCTTTGGCTTTTGTGGCCTTCTTTGA 1318
QY 450 rPheGlyMetGlyAlaIleProTyrLeuMetMetSerGluIleLeuProValSerIleLy 470
Db 1319 AGTGGGTCTGCGCCCATCCCATGGTTTCATCGTGGCTGAACCTTTCAGCCAGGGTCCACG 1378
QY 470 sSerLeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeuThrSerPheAlaIleThrMe 490
Db 1379 TCCAGCTGCCATTGCCGTTGCAGGCTTCTCCAACCTGGACCTCAAAATTTTCATTGTGGCAT 1438
QY 490 tThrThrAsnLeuMetLeuThrTrpSerValGlyGlyThrPheLeuSerTyrMetValva 510
Db 1439 GTGCTTCCAGTATGTGGAGCAACTGTGTGGTCCCTACGTCCTTCATCATCTTCACTGTGCT 1498
QY 510 lSerAlaPheThrIleValPheValValLeuTrpValProGluThrLysGly 527
Db 1499 CCTGGTTCTGTCTTTCATCTTCACTTCACTTCAAGTTCTCCTGAGACTAAAGGC 1550

RESULT 14
US-09-221-107-135
; Sequence 135, Application US/09221107
; Patent No. 6660838
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY OF LUNG CANCER
```

FILE REFERENCE: 210121.455C2
 CURRENT APPLICATION NUMBER: US/09/221.107
 CURRENT FILING DATE: 1998-12-22
 NUMBER OF SEQ ID NOS: 161
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 135
 LENGTH: 2856
 TYPE: DNA
 ORGANISM: Human
 US-09-221-107-135

Alignment Scores:
 Pred. No.: 6.64e-47 Length: 2856
 Score: 523.50 Matches: 151
 Percent Similarity: 45.91% Conservative: 96
 Best Local Similarity: 28.07% Mismatches: 229
 Query Match: 18.84% Indels: 62
 DB: 4 Gaps: 12

US-10-051-909-36 (1-553) x US-09-221-107-135 (1-2856)

QY 27 ProGlyArgProAlaSerGluLeuArgThrArgValMetGlyGlyGlySer-AsnArgG1 46
 DB 14 CCGAGTGAGCACGCCAGGAGCAGGAGACCAACGACGGGGTTCGAGTCCAG 73
 QY 46 YGlyAla-----GlyAlaGlyGluSerGlySerAspHisAspGlyValLeuAr 63
 DB 74 TGGGAGTCCCCGACCGAGCAGCAGCCTGAGCGGAGAGCGCGCTCGCACGCCGTCG 133
 QY 63 GArgProLeuLeuAsnThrGlySerTrpTyrArgMetSerSerArgGlnSerSerPheAl 83
 DB 134 CCACCGCGTACCCGG-----CGCAGCCAGAGCCACCGCAGCGCTGC 178
 QY 83 aProGlyThrSerSerMetAlaValLeuArgGluSerHisValSerAlaPheLeuCysTh 103
 DB 179 CATGGAGCCAGCAGCAAGAAGCTGACGGGTGC-----CTCATGCTGGTGT 226
 QY 103 rLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGlyPheSerSerProTh 123
 DB 227 GGGAGGAGAGTGTCTGGCTCCCTGCGAGTTGGCTACAACTGAGTCAATGCCCC 286
 QY 123 rGlnAspAlaMetValArgAspLeuAsn----- 132
 DB 287 CCAGAAGGTGATGAGGAGTTCTACACACAGACATGGTCCACCGCTATGGGAGAGCAT 346
 QY 133 -----LeuSerIleSerGluPheSerAlaPheG1 142
 DB 347 CCTGCCCCACCAAGCTCACCAGCTCTGGTCCCTCTCAGTGGCCATCTTTCT----- 398
 QY 142 ySerLeuSerAsnValGlyGlyMetValGlyAlaIleAlaSerGlyGlnMetAlaGluTy 162
 DB 399 -----GTTGGGGCATGATGGCTCCTTCTCTGTGGGCTTTTCGTTAACCG 445
 QY 162 rIleGlyArgLysGlySerLeuMetIleAlaAlaIleProAsnIleIleGlyTrpLeuAl 182
 DB 446 CTTTGGCCGCGGAATTCAATGCTGATGATGAACCTGCTGGCCCTCGTGTCCGCGGTGCT 505
 QY 182 aIleSerPheAla-----LysAspAlaSerPheLeuTyrMetGlyArgLeuLeuG1 199
 DB 506 CATGGGCTTCTCGAAACTGGGCAAGTCTTTGAGATGCTGATCCTGGGCGCTTCATCAT 565
 QY 199 uGlyPheGlyValGlyIleIleSerTyrThrValProValTyrIleAlaGluIleSerPr 219
 DB 566 CGGTGTGTACTCGGCGCTGACACAGGCTTCGTGCCCATGATGTTGGTGAAGTGTACC 625
 QY 219 oGlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIlePh 239
 DB 626 CACAGCCCTTTCGTGGGCGCTGGGACCCCTGCACCGCTGGGATCGTCGCGCATCCT 685
 QY 239 eLeuAlaTyrLeuLeuGlyMetPheIle-----ProTrpArgLeuLe 253
 DB 686 CATCGCCCGAGGTGTTGGGCTGGACTCCATCATGGGCAACAGGACCTGTGGCCCTGCT 745

QY 253 uAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPhePheIleProG1 273
 DB 746 GCTGAGCATCATCTTCCCGCCCTGCTGAGTGCAGTGCATGCTGCTGCTTCTGCCCGA 805
 QY 273 uSerProArgTrpLeuAlaLysMetAsnLeuThrGluAsp-----CysGluThrSerLe 291
 DB 806 GAGTCCCGCTTCTGCTC--ATCAACCGCAACGAGGAGAACCGGGCCAAGAGTGTGCT 862
 QY 291 uGlnValIleuArgGlyPheGluThrAspIleThrGluValAsnAspIleLysArgAl 311
 DB 863 AAAGAAGCTGCGCGG--ACAGCTGACGTGACCCATGACCTGCAGGAGATGAAGGAAGA 919
 QY 311 aValAlaSerSerLysArgThrThrIleSerPheGlnGluLeu--AsnGlnLysLy 330
 DB 920 GAGTCGGCAGATGATCGGGAGAGAGGTACCATCCTGGAGTGTTCGCTCCCCCGC 979
 QY 330 sTyrArgThrProLeuLeuGlyIleGlyLeuLeuValLeuGlnAsnLeuSerGlyI1 350
 DB 980 CTACCGCCAGCCCATCTCATCGCTGTGTGTGCTGCAGCTGTCCAGCAGCTGTCTGGCAT 1039
 QY 350 eAsnGlyValLeuPheTyrAlaSerSerIlePheLysAlaAlaGlyValThrAsnSerAs 370
 DB 1040 CAACGCTGTCTTATTACTCCACGAGCATCTTCGAGAAAGCGGGGTGCAGCAGCCTGT 1099
 QY 370 pLeuAlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThrGlyValThrTrpLe 390
 DB 1100 GTATGCCACCATGGCTCCGGTATCGTCAACACGGCCTTCACTGCTGTGCTGTTTGT 1159
 QY 390 uLeuAspArgAlaGlyArgArgIleLeuLeuIleIleSerThrSerGlyMetThrLeuCy 410
 DB 1160 GGTGGAGCGAGCAGCGCGGACCCCTGCACCTCATAGGCTCGCTGGCATGGCGGTTG 1219
 QY 410 sLeuLeuAlaValSerValPhePheLeuLysAspAsnIleSerGlnAspSerAsnSe 430
 DB 1220 TGCCATACTCATGACCATCGCGTAGCACTGCTGGAGCAGTACCTGG----- 1268
 QY 430 rTyrTyrIleLeuThrMetIleSerLeuValGlyIleValSerPheValIleThrPheSe 450
 DB 1269 -----ATGCTCTATCTGAGCATCGTGGCCATCTTTGGCTTGTGGCCTTCTTTGA 1318
 QY 450 rPheGlyMetGlyAlaIleProTrpLeuMetMetSerGluIleLeuProValSerIleLy 470
 DB 1319 AGTGGGTCTGCGCCCATCCCATGTTTCATCGTGGTGAACCTCTCAGCCAGGGTCCACG 1378
 QY 470 sSerLeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeuThrSerPheAlaIleThrMe 490
 DB 1379 TCCAGCTGCCATTCGCGTTCAGGCTTCTCCAACTGACCTCAAAATTTTATTGTGGCAT 1438
 QY 490 tThrThrAsnLeuMetLeuThrTrpSerValGlyGlyThrPheLeuSerTyrMetValVa 510
 DB 1439 GTGCTTCCAGTATGGAGCAACTGTGTGGTCCCTACGTCTTCATCATCTTCACTGTGCT 1498
 QY 510 lSerAlaPheThrIleValPheValLeuTrpValProGluThrLysGly 527
 DB 1499 CCTGGTCTCTTCTTCTTCTTCACTTCAAGTTCCTGAGACTAAAGGC 1550

RESULT 15
 US-09-466-396A-135
 ; Sequence 135, Application US/09466396A
 ; Patent No. 6696247
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Tongtong
 ; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY AND
 ; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
 ; FILE REFERENCE: 210121.455C4
 ; CURRENT APPLICATION NUMBER: US/09/466,396A
 ; CURRENT FILING DATE: 1999-12-17
 ; NUMBER OF SEQ ID NOS: 224
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 135
 ; LENGTH: 2856
 ; TYPE: DNA
 ; ORGANISM: Homo sapien

US-09-466-396A-135

Alignment Scores:

5.64e-47	Length:	2856
523.50	Matches:	151
45.91%	Conservative:	96
28.07%	Mismatches:	229
18.84%	Indels:	62
4	Gaps:	12
DB:		
Pred. No.:		
Score:		
Percent Similarity:		
Best Local Similarity:		
Query Match:		

US-10-051-909-36 (1-553) x US-09-466-396A-135 (1-2856)

QY	27	ProGlyArgProAlaSerGluLeuArgThrArgValMetGlyGlyGlySer-AsnArgG1	46
DB	14	CCGAGTGAGCACGCCAGGAGCAGGAGACCAAACGACGGGGTCCGAGTCAGAGTCGCAG	73
QY	46	YGlyAla-----GlyAlaGlyGluGluSerGlySerAspHisAspGlyValLeuAr	63
DB	74	TGGGAGTCCCGGACCGGAGCAGCAGCCTGAGCGGGAGAGCGCGCGTCCGACGCCCGTCG	133
QY	63	gArgProLeuLeuAsnThrGlySerTrpTyrArgMetSerSerArgGlnSerSerPheAl	83
DB	134	CCACCCGCGTACCCGG-----CGCAGCCAGAGCCACGACGCGCAGCGCTGC	178
QY	83	aproGlyThrSerSerMetAlaValLeuArgGluSerHisValSerAlaPheLeuCysTh	103
DB	179	CATGGAGCCCGCAGCAAGAGCTGACGGGTCCG-----CTCATGCTGGCTGT	226
QY	103	rLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGlyPheSerSerProTh	123
DB	227	GGGAGGAGCAGTGCTTGGCTCCCTGCAGTTGGCTACAACACTGGAGTCATCAATGCCCC	286
QY	123	rGlnAspAlaMetValArgAspLeuAsn-----	132
DB	287	CCAGAAGGTGATCGAGGAGTTCTACAACACGACATGGGTCCACCGCTATGGGAGAGCAT	346
QY	133	-----LeuSerIleSerGluPheSerAlaPheG1	142
DB	347	CCTGCCCCACCGCTCACACCGCTCTGGTCCCTCTCAGTGGCCATCTTTCT-----	398
QY	142	ySerLeuSerAsnValGlyGlyMetValGlyAlaIleAlaSerGlyGlnMetAlaGluTy	162
DB	399	-----GTTGGGGGCATGATTGGCTCCTTCTCTGTGGGCCCTTTTCGTTAACCG	445
QY	162	rIleGlyArgLysGlySerLeuMetIleAlaAlaIleProAsnIleIleGlyTrpLeuAl	182
DB	446	CTTTGGCCGGCGGAATTCAATGCTGATGATGAACCTGCTGGCCTTCGTGTCCGCCGTGCT	505
QY	182	alleSerPheAla-----LysAspAlaSerPheLeuTyrMetGlyArgLeuLeuG1	199
DB	506	CATGGGCTTCTCGAAACTGGCAAGTCCCTTTGAGATGCTGATCTCTGGGCCGCTTCATCAT	565
QY	199	uGlyPheGlyValGlyIleIleSerTyrThrValProValTyrIleAlaGluIleSerPr	219
DB	566	CGGTGTGTACTGCGGCCTGACCACAGGCTTCGTGCCCATGTATGTGGGTGAAGTGTCAAC	625
QY	219	oGlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIlePh	239
DB	626	CACAGCCTTTCGTGGGGCCCTGGGCACCCCTGCACACAGCTGGGCATCGTCTCGGCATCCT	685
QY	239	eLeuAlaTyrLeuLeuGlyMetPheIle-----ProTrpArgLeuLe	253
DB	686	CATGCCCCAGGTGTTCCGGCTGGACTCCATCATGGGCAACAAGGACCTGTGGCCCTGCT	745
QY	253	uAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPhePheIleProG1	273
DB	746	GCTGAGCATCATCTTCATCCCGGCCCTGCTGGAGTGCATCGTGTGCCCTTCTGCCCGGA	805
QY	273	userProArgTrpLeuAlaLysMetAsnLeuThrGluAsp-----CysGluThrSerLe	291
DB	806	GAGTCCCGCTTCTGTCTC---ATCAACCGCAACGAGGAGAACCGGGCCCAAGAGTGTCT	862
QY	291	uGlnValLeuArgGlyPheGluThrAspIleThrGluValAsnAspIleLysArgAl	311

Db	863	AAAGAAAGCTGGCGGG---ACAGCTGACGTGACCCATGACCTGCAGGAGATGAAGGAAGA	919
Qy	311	avalAlaSerSerSerLysArgThrThrIleSerPheGlnGluLeu--AsnGlnLysLy	330
Db	920	GATCGGCAGATGATCGGGAGAAAGGTACCATCTCTGGAGCTGTTCCTCGCTCCCGCGC	979
Qy	330	sTyrArgThrProLeuLeuLeuGlyIleGlyLeuLeuValLeuGlnAsnLeuSerGlyI1	350
Db	980	CTACCGCCAGCCCATCTCTATCGTGTGTGTCGTGCAGCTGTCCCAGCAGCTGTCTGGCAT	1039
Qy	350	eAsnGlyValLeuPheTyrAlaSerSerIlePheLysAlaAlaGlyValThrAsnSerAs	370
Db	1040	CAACGCTGTCTTCTATTACTCCACGAGCATCTTCGAGAAGCGGGGTGCAGCAGCCTGT	1099
Qy	370	pLeuAlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThrGlyValThrTrpLe	390
Db	1100	GATGCCACCATTTGGCTCCGGTATCGTCAACACGGCCTCACTGTCTGTCGTGTTGT	1159
Qy	390	uLeuAspArgAlaGlyArgArgIleLeuLeuIleIleSerThrSerGlyMetThrLeuCy	410
Db	1160	GGTGGAGCGAGCAGGCCGCGGACCCTGCACCTCATAGGCCTCGCTGGCATGGCGGTTG	1219
Qy	410	sLeuLeuAlaValSerValValPhePheLeuLysAspAsnIleSerGlnAspSerAsnSe	430
Db	1220	TGCCATATCATGACCATCGCGCTAGCACTGTCTGGAGCAGCTACCTGG-----	1268
Qy	430	rTyrTyrIleLeuThrMetIleSerLeuValGlyIleValSerPheValIleThrPheSe	450
Db	1269	-----ATGTCTTATCTAGCATCGTGGCCATCTTTGGCTTTTGTGGCCCTCTCTTGA	1318
Qy	450	rPheGlyMetGlyAlaIleProTrpLeuMetMetSerGluIleLeuProValSerIleLy	470
Db	1319	AGTGGTCTCTGGCCCCATCCCATGGTTTCATCGTGGCTGAACCTTTTCAGCCAGGGTCCACG	1378
Qy	470	sSerLeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeuThrSerPheAlaIleThrMe	490
Db	1379	TCCAGCTGCCATTGCCGTTGCAGGCTTCTCCAACCTGGACCTCAAATTTTCATTGTGGGCAT	1438
Qy	490	tThrThrAsnLeuMetLeuThrTrpSerValGlyGlyThrPheLeuSerTyrMetValVa	510
Db	1439	GTGCTTCCAGTATGTGGAGCAACTGTGTGGTCCCTACGCTTTCATCATCTTTCACCTGTGCT	1498
Qy	510	lSerAlaPheThrIleValPheValValLeuTrpValProGluThrLysGly	527
Db	1499	CCTGGTTCTGTTCTTCTCATCTTCACCTACTTCAAAGTTCTCTGAGACTAAAGGC	1550

Search completed: October 13, 2004, 23:31:06
Job time : 176.894 secs

is Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 13, 2004, 16:55:37 ; Search time 788.832 Seconds
(without alignments)
3566.233 Million cell updates/sec

Title: US-10-051-909-36
Perfect score: 2779
Sequence: 1 PSSSSFRPAGKKKKKNGG.....TIFVLSIQRLQWLPECLS 553

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3372238 seqs, 2543544081 residues

Total number of hits satisfying chosen parameters: 6744476

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO spool h/US10051909/runat_13102004_123339_20010/app_query.fasta_1.1678
-DB=Published Applications NA -QFMT=fastap -SUFFIX=rpnb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORV=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10051909@cgn 1 1 953 @runat_13102004_123339_20010
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result *
Query

SUMMARIES

No.	Score	Match	Length	DB	ID	Description
1	2777	99.9	1953	13	US-10-051-909-35	Sequence 35, Appl
2	2133	76.8	2003	17	US-10-437-963-61061	Sequence 61061, A
3	2129.5	76.6	1797	16	US-10-425-114-26596	Sequence 26596, A
4	2104	75.7	1875	16	US-10-425-114-26550	Sequence 26550, A
5	1806.5	65.0	2049	16	US-10-425-114-24968	Sequence 24968, A
6	1718	61.8	2202	16	US-10-424-599-132422	Sequence 132422, A
7	1374	49.4	1304	17	US-10-437-963-63465	Sequence 63465, A
8	1221.5	44.0	957	16	US-10-425-114-28778	Sequence 28778, A
9	1185	42.6	1283	16	US-10-424-599-131066	Sequence 131066, A
10	1182.5	42.6	1077	17	US-10-767-795-565	Sequence 565, App
11	1139.5	41.0	2105	16	US-10-424-599-132426	Sequence 132426, A
12	1139	41.0	1027	16	US-10-425-114-5566	Sequence 5566, App
13	946.5	34.1	990	17	US-10-767-795-566	Sequence 566, App
14	920.5	33.1	1473	9	US-09-938-842A-2331	Sequence 2331, Ap
15	920.5	33.1	1473	11	US-09-938-842A-2331	Sequence 2331, Ap
16	909.5	32.7	1785	17	US-10-437-963-64291	Sequence 64291, A
17	899	32.3	1359	9	US-09-938-842A-1627	Sequence 1627, Ap
18	899	32.3	1359	11	US-09-938-842A-1627	Sequence 1627, Ap
19	893	32.1	1768	15	US-10-310-154-357	Sequence 357, App
20	877.5	31.6	1407	9	US-09-938-842A-1291	Sequence 1291, Ap
21	877.5	31.6	1407	11	US-09-938-842A-1291	Sequence 1291, Ap
22	777	28.0	1374	9	US-09-938-842A-1020	Sequence 1020, Ap
23	777	28.0	1374	11	US-09-938-842A-1020	Sequence 1020, Ap
24	759	27.3	1635	17	US-10-437-963-4124	Sequence 4124, Ap
25	733	26.4	1231	16	US-10-424-599-132423	Sequence 132423, A
26	693.5	25.0	1779	16	US-10-424-599-104365	Sequence 104365, A
27	679.5	24.5	778	16	US-10-425-114-22347	Sequence 22347, A
28	669.5	24.1	2080	15	US-10-168-651-28	Sequence 28, Appl
29	666	24.0	1258	16	US-10-424-599-132425	Sequence 132425, A
30	663	23.9	1473	9	US-09-938-842A-1669	Sequence 1669, Ap
31	663	23.9	1473	11	US-09-938-842A-1669	Sequence 1669, Ap
32	658.5	23.7	1862	15	US-10-115-831-137	Sequence 137, App
33	657.5	23.7	600	17	US-10-021-323-14004	Sequence 14004, A
34	625.5	22.5	890	16	US-10-425-114-127	Sequence 127, App
35	623.5	22.4	2487	15	US-10-172-118-1685	Sequence 1685, Ap
36	623.5	22.4	2487	16	US-10-342-887-1685	Sequence 1685, Ap
37	623.5	22.4	2487	17	US-10-755-889-109	Sequence 109, App
38	621	22.3	2356	15	US-10-157-031-358	Sequence 358, App
39	621	22.3	2696	10	US-09-814-353-19608	Sequence 19608, A
40	617	22.2	501	16	US-10-425-114-34725	Sequence 34725, A
41	606	21.8	1905	9	US-09-734-569-171	Sequence 171, App
42	605	21.8	792	16	US-10-425-114-5672	Sequence 5672, Ap
43	604	21.7	460	17	US-10-767-701-15010	Sequence 15010, A
44	602	21.7	1405	17	US-10-767-701-12467	Sequence 12467, A
45	598.5	21.5	475	17	US-10-767-701-18449	Sequence 18449, A

ALIGNMENTS

RESULT 1
US-10-051-909-35
; Sequence 35, Application US/10051909
; Publication No. US20020199217A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Helentjaris, Tim
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: B01163 US CIP
; CURRENT APPLICATION NUMBER: US/10/051,909
; PRIOR FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 60/083,044
; PRIOR FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 35
; LENGTH: 1953
; TYPE: DNA
; ORGANISM: Zea mays

NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 61061
LENGTH: 2003
TYPE: DNA
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_62530C.1
US-10-437-963-61061

Alignment Scores:

Pred. No.: 2, 42e-203 Length: 2003
Score: 2133.00 Matches: 434
Percent Similarity: 87.77% Conservative: 47
Best Local Similarity: 79.20% Mismatches: 51
Query Match: 76.75% Indels: 17
DB: 17 Gaps: 7

US-10-051-909-36 (1-553) x US-10-437-963-61061 (1-2003)

QY 1 ProSerSerSerSerSerPheArgProAlaGlyLysLysLysLysLysAsnGlnGly 20
Db 55 CCTTCCAATTCACT-----CGTCTCTCTCTCGTAACTCG----- 90
QY 21 LeuArgArgGluAlaValProGlyArgProAlaSerGluLeuArgThrArg-----Val 38
Db 91 ---TAAAGAGATCTCGTCTCGCTCGGGAATCCAAACCACTTGCCTCGGGAAGTAGAGATCGG 147
QY 39 MetGlyGlyGlySerAsnArg--GlyGlyAlaGlyAlaGlyGluGlySerGlySerAsp 57
Db 148 GAGGGCGGCCCATGAAAGGGCGGGCGGCGTGGCGGCGATGAGAGCGGAAGCGAC 207
QY 58 HisAsp-----GlyValLeuArgArgProLeuLeu--AsnThrGlySerTrpTyrArg 74
Db 208 TACGAGAGCGGGCGGCGGATCGGAAGCCGCTGTGATGCACACGGGGAGTGGTACAGG 267
QY 75 MetSerSerArgGlnSerSerPhe--AlaProGlyThrSerSerMetAlaValLeuArg 93
Db 268 ATGGGGTGGCGGCGGAGGAGCTCACCGGGCGGGGACCTCGTCCATGGCCATCTACGC 327
QY 94 GluSerHisValSerAlaPheLeuCysThrLeuLeuValAlaLeuGlyProIleGlnPhe 113
Db 328 GAGTCCCACGTCTCCGCTCTCTGACGCTCATCGTGGCGCTCGGCCCCATCCAGTTC 387
QY 114 GlyPheThrSerGlyPheSerProThrGlnAspAlaMetValArgAspLeuAsnLeu 133
Db 388 GGATTACGGGGGGCTTCTCTCCCGACGCGAGAGCCATCATCCGAGACCTCGACCTC 447
QY 134 SerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyGlyMetValGlyAla 153
Db 448 ACCCTCTCCGAGTCTCGGTGTTCCGATCGCTGTCAACGTCGGCGCCATGGTGGGGCG 507
QY 154 IleAlaSerGlyGlnMetAlaGluTyrIleGlyArgLysGlySerLeuMetIleAlaAla 173
Db 508 ATTGCCAGTGGTTCAGATGGCGGAGTACATTGGGCGCAAGGGTCAATTGATGATGCTGCA 567
QY 174 IleProAsnIleIleGlyTrpLeuAlaIleSerPheAlaLysAspAlaSerPheLeuTyr 193
Db 568 ATTCCAAACATCATTTGGTGGCTTGGCATCTCTCTTGCAGAAAGACTCATCGTTCTTTAT 627
QY 194 MetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTyrThrValProValTyr 213
Db 628 ATGGGACGATTGCTCGAGGGGTTGGTGTGGTGTGTCATCTCTTATACGGTGGCAGTTTAC 687
QY 214 IleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSer 233
Db 688 ATAGCAGAAATATCACCTCAAAACATGAGAGGTGCTCTTGGCTCAGTGAATCAGTTATCT 747
QY 234 ValThrPheGlyIlePheLeuAlaTyrLeuLeuGlyMetPheIleProTrpArgLeuLeu 253
Db 748 GTAACCGTTGGTATATTGGCATATTGTCTCGGCATGTTTGTCTCTTGGAGGCTTCT 807
QY 254 AlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPhePheIleProGlu 273

Db 808 CTTGTAATAGGAATCTTGCTTGCCTTGCACTGTGTGTGATACCTGGCCTATTCTTCATTCCAGAA 867
QY 274 SerProArgTrpLeuAlaLysMetAsnLeuThrGluAspCysGluThrSerLeuGlnVal 293
Db 868 TCCCCAAGATGGTTGGCAAAGATGAACATGATGATGATTTTGAGACITCTTTTACAAGTT 927
QY 294 LeuArgGlyPheGluThrAspIleThrThrGluValAsnAspIleLysArgAlaValAla 313
Db 928 CTGAGGGGATTTGAGACTGACATCAGCGCGGAAGTGAATGATATAAGAGAGAGTAGCG 987
QY 314 SerSerSerLysArgThrThrIleSerPheGlnGluLeuAsnGlnLysLysTyrArgThr 333
Db 988 TCAGCAAAACAAAGGACACGATCCGTTTTCAGAATAATAACCAAGAAATACCGCACA 1047
QY 334 ProLeuLeuGlyIleGlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyVal 353
Db 1048 CCCCTAATACTAGGAATTGGCTTACTTGTACTGCAACAGCTAAGTGAATCAATGGAATA 1107
QY 354 LeuPheTyrAlaSerSerIlePheLysAlaAlaGlyValThrAsnSerAspLeuAlaThr 373
Db 1108 TTGTTTATGCAGGTAGCATCTTCAAAGCAGCAGGCTCTCAGAAACAGTGACTTGGCTACA 1167
QY 374 CysSerLeuGlyAlaIleGlnValLeuAlaThrGlyValThrThrTrpLeuLeuAspArg 393
Db 1168 TGTGCACTTGGTGTATCCAGGTCTTGTCTACAGGAGTTACAACCTGGTTATTAGACAGA 1227
QY 394 AlaGlyArgArgIleLeuLeuIleIleSerThrSerGlyMetThrLeuCysLeuLeuAla 413
Db 1228 GCTGCCGACGGATCTCTCTTATCATCTCTTCTGCTGGATGACTCTAAGCCTCTCTGCA 1287
QY 414 ValSerValValPhePheLeuLysAspAsnIleSerGlnAspSerAsnSerTyrTyrIle 433
Db 1288 GTTGTGTGTATTTTCTCAAGGATAGCATTTTCAAGATTTCTACATGTACTACACC 1347
QY 434 LeuThrMetIleSerLeuValGlyIleValSerPheValIleThrPheSerPheGlyMet 453
Db 1348 TTAAGTATGATCTCTCTTGGTGTCTTGTGGCTTTTGTAAATCGCCTTCTCCTTCGGTATG 1407
QY 454 GlyAlaIleProTrpLeuMetMetSerGluIleLeuProValSerIleLysSerLeuGly 473
Db 1408 GGTGCCATTCCATGGATCATATGTTCAGAGATCCTCCCGTTAGTATCAAGAGTCTCGCA 1467
QY 474 GlySerIleAlaThrLeuAlaAsnTrpLeuThrSerPheAlaIleThrMetThrThrAsn 493
Db 1468 GGAAGCTTTCGACGCTCGCAACTGGCTTACATCTTGGATTAACAATGACAGCAAC 1527
QY 494 LeuMetLeuThrTrpSerValGlyGlyThrPheLeuSerTyrMetValValSerAlaPhe 513
Db 1528 TTGATGCTTAGCTGGAGTGTGGAGGGACCTTTGTGTCTACATGGTCTGTGATGCTTTC 1587
QY 514 ThrIleValPheValLeuTrpValProGluThrLysGly***AsnSerArgGlyAsp 533
Db 1588 ACCCTCGTGTTCGTATCTCTTGGTGGCAGAGACAAAGGAAG-AACTCTCGAAGAGAT 1646
QY 534 ThrIlePheValSerLeuSerIle 541
Db 1647 ACAATGGTCTTCCGCTGAGCCTT 1670

RESULT 3

US-10-425-114-26596
; Sequence 26596, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114

FEATURE:
OTHER INFORMATION: Clone ID: LIB3689-227-G11_FLI
US-10-425-114-24968

Alignment Scores:

Pred. No.: 13e-170 Length: 2049
Score: 1806.50 Matches: 366
Percent Similarity: 81.04% Conservatives: 70
Best Local Similarity: 68.03% Mismatches: 79
Query Match: 65.01% Indels: 26
DB: 16 Gaps: 4

US-10-051-909-36 (1-553) x US-10-425-114-24968 (1-2049)

QY 34 LeuArgThrArgValMetGlyGlySerAsnArgGlyGlyAlaGlyAlaGlyGluGlu 53
Db 314 CTCAGGCATGAGCTCCGGACCGAGAGTGGCGGGAGGATG-GGGGCAGGACG-TCC 371
QY 54 SerGlySerAspHisAspGlyValLeuArgArgProLeuLeuAsnThrGlySerTrpTyr 73
Db 372 TCCGCCTCCGAC-----CTGCGAAGCCGCTCATCAACACCGGGAGCTGGTAC 419
QY 74 Arg-----MetSerSerArgGlnSerSerPheAlaProGlyThr 86
Db 420 CGCATGCCGCGCGGGTGGCGTGTATGGGCTCGCGGCAATCCAGCTCATGGAGCGCTTG 479
QY 87 SerSerMetAlaVal---LeuArgGluSerHisValSerAlaPheLeuCysThrLeuIle 105
Db 480 GGCTCCTCTGCGTCTCTCCGCGACGTCGCTATCTCGGCCACGCTCGCACGCTTATT 539
QY 106 ValAlaLeuGlyProIleGlnPheGlyPheThrSerGlyPheSerProThrGlnAsp 125
Db 540 GTCGCGCTAGGTCCTCATCCAGTTCGGTTTCACATGCGGCTACTCTCGGCCACGCGAGGAC 599
QY 126 AlaMetValArgAspLeuAsnLeuSerIleSerGluPheSerAlaPheGlySerLeuSer 145
Db 600 GCCATCATGTGATCTCGGCTCTCCCTCTCTGAGTTCTCCCTCTCTCGGTTTCAATTATCT 659
QY 146 AsnValGlyGlyMetValGlyAlaIleAlaSerGlyGlnMetAlaGluTyrIleGlyArg 165
Db 660 AATGTAGGGCGAGTGGTAGCGCCATCTCCAGTGGGCAACTTGCAGAGTATATCGGCCGC 719
QY 166 LysGlySerLeuMetIleAlaAlaIleProAsnIleIleGlyTrpLeuAlaIleSerPhe 185
Db 720 AAGGGGTCTCTCATGATCGCTCGGATTCCAAACATAATTGGGTGGCTCGCGATATCATTC 779
QY 186 AlaLysAspAlaSerPheLeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIle 205
Db 780 GCAAAAGATTCCCTCTTCTTGTATTGGTCTCGGCTGCTAGAGGATTTGGAGTCCGGTGA 839
QY 206 IleSerTyrThrValProValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAla 225
Db 840 ATATCGTATACAGTACCGGTTTATATTGAGAAATCGCTCCTCAGGATCAGAGGGAGGT 899
QY 226 LeuGlySerValAsnGlnLeuSerValThrPheGlyIlePheLeuAlaTyrLeuLeuGly 245
Db 900 CTTGGTTCTGTCAATCAGCTCTCCGTCACGATTGGTATATTGCTTGCCTACTTGTGGC 959
QY 246 MetPheIleProTrpArgLeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIle 265
Db 960 ATGTTTGTTCCTGGAGAAATCTTGTCTTCTAGGCATTTTACCTTGTTCATTCCTGATT 1019
QY 266 ProGlyLeuPhePheIleProGluSerProArgTrpLeuAlaLysMetAsnLeuThrGlu 285
Db 1020 CCTGGACTGTCTTGTGCTGAAATCCCAAGGTGGTTGGCAAAATGGGGAAGATGGAG 1079
QY 286 AspCysGluThrSerLeuGlnValLeuArgGlyPheGluThrAspIleThrThrGluVal 305
Db 1080 GATTTTGAATATTCTTCAAGTTCTGAGAGGATTTTCAGACAGATATCACAGCAGAGTA 1139
QY 306 AsnAspIleLysArgAlaValAlaSerSerLysArgThrThrIleSerPheGlnGlu 325
Db 1140 AATGAAATAAGAGATCATTAGCATCATCGAGGAGGAGGACCAACCATTAAGGTTTCGCTCAT 1199

QY 326 LeuAsnGlnLysLysTyrArgThrProLeuLeuLeuGlyIleGlyLeuLeuValLeuGln 345
Db 1200 ATCAAAACAGAGAGATACAGTGTTCCTCCCTTGTGATAGGAATCGGTCTCTCTGCTGCAG 1259
QY 346 AsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSerIlePheLysAlaAlaGly 365
Db 1260 CAGCTAAGTGGTGTCAATGGCATTCTATTTATGCTCGGAGCATCTTCAAAGCTGCTGGT 1319
QY 366 ValThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThrGly 385
Db 1320 ATTACAAACAGTATCTAGCAACATTTGGTTTGGGGCTGTTTCAGGTGATTGCTACTGGA 1379
QY 386 ValThrThrTrpLeuLeuAspArgAlaGlyArgArgIleLeuLeuIleSerThrSer 405
Db 1380 GTGACAAACCTGGTTGACTGACAAAGCTGGTGAAGGCTTCTTCTCATATTATTCACACCA 1439
QY 406 GlyMetThrLeuCysLeuLeuAlaValSerValPhePheLeuLysAspAsnIleSer 425
Db 1440 GGAATGGTCACTTACTTCTTATTGTTTCTGTGTCTATTTTGTGAAGGACAACATAGCT 1499
QY 426 GlnAspSerAsnSerTyrTyrIleLeuThrMetIleSerLeuValGlyIleValSerPhe 445
Db 1500 GCTGTTTCGCACTTACTCTGTAATAGTATGCTTTTCACTGGCTGGACTTGTGGCATTT 1559
QY 446 ValIleThrPheSerPheGlyMetGlyAlaIleProTrpLeuMetSerGluIleLeu 465
Db 1560 GTGATTGCATTTCTCTGGCTTGGGAGCGATTCCGTGGATCATATTATGCTGAGATCCTT 1619
QY 466 ProValSerIleLysSerLeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeuThrSer 485
Db 1620 CCTGTTAACATCAAGAGCCTTGTGGAAGTGTTCGACCCCTGGCGAGTGGAGAACATTGCT 1739
QY 486 PheAlaIleThrMetThrThrAsnLeuMetLeuThrTrpSerValGlyThrPheLeu 505
Db 1680 TGGGCCATTACATGACGGCAAGCCTGATGTTGAACCTGGAGCAGTGGAGAACATTGCT 1739
QY 506 SerTyrMetValValSerAlaPheThrIleValPheValValLeuTrpValProGluThr 525
Db 1740 ATCTACGCCGCTGCTGTCTACCATGGCCCTCATTTTCGTGTGCTTGTGGTGCCTGAGACC 1799
QY 526 LysGly***AsnSerArgGlyAspThrIlePheValSerLeuSerIleGlnArgGlnLeu 545
Db 1800 AAGGGAAG-AACGCTAGAGGAAATCGCCTTCTCATTCGCTGACACGCTCGTCATGATCTA 1858
QY 546 -GlnTrp-----LeuProGluCysLeu 552
Db 1859 GGTATGGGAAGCCACACACCATGTCCTTGTGATCTGCCCGCGTGCCTT 1910

RESULT 6

US-10-424-599-132422
; Sequence 132422, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 132422
; LENGTH: 2202
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(2202)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:

OTHER INFORMATION: Clone ID: PAT_MRT3847_90584C.1
US-10-424-599-132422

Alignment Scores:

Pred. No.: 1,07e-161 Length: - 2202
Score: 1718.00 Matches: 334
Percent Similarity: 83.67% Conservative: 76
Best Local Similarity: 68.16% Mismatches: 76
Query Match: 61.82% Indels: 5
DB: 16 Gaps: 2

US-10-051-909-36 (1-553) x US-10-424-599-132422 (1-2202)

QY	52	GluGluSerGlySerAspHisAspGlyValLeuArgProLeuLeuAsnThrGlySer	71
DB	255	GAAGAGAGTGGCGATCGCAGGGAT-----CTTCAGAAACCGTTCTCCACACAGGGAGT	308
QY	72	TrpTyrArgMetSerSerArgGlnSerSerPheAlaProGlyThrSerSerMetAlaVal	91
DB	309	TGGTACAAGATGGTTCCAGGCAGTCCAGCATATG-----GGATCCTCCACTCATGTT	362
QY	92	LeuArgGluSerHisValSerAlaPheLeuCysThrLeuIleValAlaLeuGlyProIle	111
DB	363	ATCGCGACGGCGCGTTTCGGTCCCTCTTCTCGCTCCTCATCGTCGCTTGGGTCCCAT	422
QY	112	GlnPheGlyPheThrSerGlyPheSerSerProThrGlnAspAlaMetValArgAspLeu	131
DB	423	CAATTCGGCTTCACGTGTGGTATTCTTCTCCAAACCAAGGGGCTATAGTTCGGATCTA	482
QY	132	AsnLeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyGlyMetVal	151
DB	483	AACCTCTCTATTTCGGAGTTTCTTCTTGGATCTTGTCTAATGTGGGAGCAATGGTG	542
QY	152	GlyAlaIleAlaSerGlyGlnMetAlaGluTyrIleGlyArgLysGlySerLeuMetIle	171
DB	543	GGAGCTATAGCTAGTGTGCAGATAGCTGAATACATCGGGCGCAAGGGTCAATGATG	602
QY	172	AlaAlaIleProAsnIleGlyTrpLeuAlaIleSerPheAlaLysAspAlaSerPhe	191
DB	603	GCTCGCATCCCCAATATAATAGGGTGGCTTGCTATTCTTTTGGCCAAAGATCCTCGTT	662
QY	192	LeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTyrThrValPro	211
DB	663	TTGTATATGGGAGGTTGTTCGAAGGTTTGGCGTTGGGATTAATCTCTATATGGTGCCT	722
QY	212	ValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGln	231
DB	723	GTTTATATAGCTGAGATTGCACCTCAAACTTGAGAGTGGCCTTGGGTCACTGAACCCAG	782
QY	232	LeuSerValThrPheGlyIlePheLeuAlaTyrLeuLeuGlyMetPheIleProTrpArg	251
DB	783	CTCTCTGTTACAATTGGCAATTATGCTGGCTTATCTGTGGGTCTTTTGTCAACTGGAGA	842
QY	252	LeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPhePheIle	271
DB	843	GTCCTTGCAATTCTAGGAATTTGCGCTTGTACAGTATTAATACCTGGATTATTTTCATA	902
QY	272	ProGluSerProArgTrpLeuAlaLysMetAsnLeuThrGluAspCysGluThrSerLeu	291
DB	903	CCTGAATCCCCCAGATGGTTGGCCAAGATGGGGATGATAGATGAGTTTGAGACTTCTTTG	962
QY	292	GlnValLeuArgGlyPheGluThrAspIleThrThrGluValAsnAspIleLysArgAla	311
DB	963	CAAGTGTACGAGGATTTGACACTGATATATCTGTGTGAAGTACATGAAATTAAGAGATCT	1022
QY	312	ValAlaSerSerLysArgThrThrIleSerPheGlnGluLeuAsnGlnLysLysTyr	331
DB	1023	GTGGCTTCAACGGGAAAAGAGCTGCAATCCGATTTCAGATCTCAAGACGAAAGATAT	1082
QY	332	ArgThrProLeuLeuGlyIleGlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsn	351
DB	1083	TGGTCCCGTTAATGGTTGGTATGGATTACTTGTTCCTCAGCAATTATCTGGTATCAAT	1142

QY	352	GlyValLeuPheTyrAlaSerSerIlePheLysAlaAlaGlyValThrAsnSerAspLeu	371
DB	1143	GGAAATTTGTTCTATTCACTACCATCTTTGCAAAATGCAGGAATTTTCATCCAGCAAGCT	1202
QY	372	AlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThrGlyValThrTrpLeuLeu	391
DB	1203	GCTACAGTTGGACTTGGAGCCGTTCCAGGTTCATAGCAACTGGAATTTCCACATGGTTGGTG	1262
QY	392	AspArgAlaGlyArgArgIleLeuLeuIleIleSerThrSerGlyMetThrLeuCysLeu	411
DB	1263	GACAAAAGTGGCCGAGGCTGCTTAATAATATCTCTATCTCTATGTAATGACAGTTAGCCTT	1322
QY	412	LeuAlaValSerValValPhePheLeuLysAspAsnIleSerGlnAspSerAsnSerTyr	431
DB	1323	CTCATTTGTTCTATAGCATTTTATCTGGAGGGGTTGTATCAGAGGATTTCACATTTATTC	1382
QY	432	TyrIleLeuThrMetIleSerLeuValGlyIleValSerPheValIleThrPheSerPhe	451
DB	1383	AGCAATTTGGGAATAGTTTCTATTGTTGGACTCGTGGCTATGGTATGGGTTCTCTCTA	1442
QY	452	GlyMetGlyAlaIleProTrpLeuMetMetSerGluIleLeuProValSerIleLysSer	471
DB	1443	GGTCTGGGACCCATCCCTTGGCTTATAATGTCTGAGATACTTCCAGTGAATATAAGGCG	1502
QY	472	LeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeuThrSerPheAlaIleThrMetThr	491
DB	1503	CTTGTCTGGCAGCATAGCGACAATGGGAAATTGGCTGATTTCGTGGGGATCACGATGACT	1562
QY	492	ThrAsnLeuMetLeuThrTrpSerValGlyGlyThrPheLeuSerTyrMetValValSer	511
DB	1563	GCTAACTTGCTTTTGAAATTGGAGCAGTGGAGGGACATTTACAATCTACACAGTCGTAGCT	1622
QY	512	AlaPheThrIleValPheValValLeuTrpValProGluThrLysGly***AsnSerArg	531
DB	1623	GCCTTTACTATAGCTTTTATAGCAATGTGGTTCCTGAGACCAAGGAAG-AACATTGGA	1681
QY	532	GlyAspThrIlePheValSerLeuSerIle	541
DB	1682	AGAAATTCAGTTTCTTCATCAGATAGATATA	1711

RESULT 7

US-10-437-963-63465
; Sequence 63465, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 63465
; LENGTH: 1304
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_64704C.1

US-10-437-963-63465	Alignment Scores:	1,45e-127	Length:	1304
	Pred. No.:	1374.00	Matches:	293
	Score:	68.89%	Conservative:	48
	Percent Similarity:	59.19%	Mismatches:	58
	Best Local Similarity:	49.44%	Indels:	97
	Query Match:			


```
QY 481 AsnTrpLeuThrSerPheAlaIleThrMetThrThrAsnLeuMetLeuThrTrpSerVal 500
Db 842 AACTGGTTGTTTCCCTGGTTGTTACATTGACAGCAATATGCTTTGGATTGGAGCTCT 901
QY 501 GlyGlyThrPheLeuSerTyrMetValValSerAlaPheThrIleValPheValValLeu 520
Db 902 GGAGGAACCTTCACCATATATGCAGTTGATGCTTGCCTTGACTGTAGTATTGTTACCACT 961
QY 521 TrpValProGluThrLysGly***AsnSerArgGlyAspThrIlePheValSerLeuSer 540
Db 962 TGGGTCCCTGAGACAAAG-----GGAAAACTATT----- 991
QY 541 IleGlnArgGlnLeuGlnTrp 547
Db 992 -----GAAGAAATCCAATGG 1006

RESULT 10
US-10-767-795-565
; Sequence 565, Application US/10767795
; Publication No. US20040181830A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Cao, Yongwei
; APPLICANT: Zhou, Yihua
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53534)B
; CURRENT APPLICATION NUMBER: US/10/767,795
; CURRENT FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 117596
; SEQ ID NO 565
; LENGTH: 1077
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: GOSHI-09MAY01-C14353_2
US-10-767-795-565

Alignment Scores:
Pred. No.: 1.68e-108 Length: 1077
Score: 1182.50 Matches: 230
Percent Similarity: 76.58% Conservative: 61
Best Local Similarity: 60.53% Mismatches: 55
Query Match: 42.55% Indels: 34
DB: 17 Gaps: 2

US-10-051-909-36 (1-553) x US-10-767-795-565 (1-1077)
QY 12 LysLysLysLysLysLysAsnGlnGlyLeuArgArgGluAlaValProGlyArgProAla 31
Db 37 AGAAAAAAGAGGAGGAAACATGATTTTCAGGATGATTTCTGAAGATGGGAGGAC--- 93
QY 32 SerGluLeuArgThrArgValMetGlyGlySerAsnArgGlyGlyAlaGlyAlaGly 51
Db 93 ----- 93
QY 52 GluGluSerGlySerAspHisAspGlyValLeuArgArgProLeuLeuAsnThrGlySer 71
Db 94 -----GTACGGAAGCCATTTCTGCATACCGGAAGC 123
QY 72 TrpTyrArgMetSerSerArgGlnSerSerPheAlaProGlyThrSerSerMetAlaVal 91
Db 124 TGGTATCGAATGGTTCCAGGACGGGATCCAGTATGCTGGGCTCATCTCAA-----GTT 177
QY 92 LeuArgGluSerHisValSerAlaPheLeuLeuCysThrLeuIleValAlaLeuGlyProIle 111
Db 178 ATTCGTGATAGTTCCATCTCTGTTGTAGCTTGCCTTTGATTTGTTGCTTGGTCTCATC 237
QY 112 GlnPheGlyPheThrSerGlyPheSerSerProThrGlnAspAlaMetValArgAspLeu 131
Db 238 CAATTGGTTTCACTAATGCTACTCTTCTCTACACAATCTGCAATCATCAAGGAACCT 297
QY 132 AsnLeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyGlyMetVal 151
```

```
Db 298 GGGTTAACTGTATCAGAGTATTCTGTATTGGTTCTTTATCAATGTGGTGCCATGATT 357
QY 152 GlyAlaIleAlaSerGlyGlnMetAlaGluTyrIleGlyArgLysGlySerLeuMetIle 171
Db 358 GGAGCAGTAGTTAGTGGTCAGATGGCAGAGTATATAGGGCGAAAGGGTCTTTGATGATT 417
QY 172 AlaAlaIleProAsnIleIleGlyTrpLeuAlaIleSerPheAlaLysAspAlaSerPhe 191
Db 418 GCTGCTATTCCATAATATATTGGATGGCTTCTATATCTTTTGCAATGACGTATCTTCC 477
QY 192 LeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTyrThrValPro 211
Db 478 CGGTTCCGTCAGGCTGCTGGAAGGTTTCGGTGTGGGAATAATCTCTTACACGGTGCCT 537
QY 212 ValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGln 231
Db 538 GTGTATATAGCTGAGATAGCACCCGAAACTTGAGAGGCAGCTTGGGTCTCTGTGAATCAG 597
QY 232 LeuSerValThrPheGlyIlePheLeuAlaTyrLeuLeuGlyMetPheIleProTrpArg 251
Db 598 TTGTCTGTTACAACTGGAACAATGGTAGCTATTGCTGGACTTTTGTGGGGTGGAGG 657
QY 252 LeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPheIle 271
Db 658 ATACTTGCAATTTTAGGAGGTGTTGCTTGTAACAATTTGATTCAGGCTATTTTTCATT 717
QY 272 ProGluSerProArgTrpLeuAlaLysMetAsnLeuThrGluAspCysGluThrSerLeu 291
Db 718 CCAGAACTCTCCTAGATGGCTGGCAAAATGGAATGACCGAAGATTTCCGAAGCTTCTCT 777
QY 292 GlnValLeuArgGlyPheGluThrAspIleThrThrGluValAsnAspIleLysArgAla 311
Db 778 CAAGTTCTTAGGGGTTTCGATACCGATATTTCTATCGAAGTCAATGAATCAAGAGGTCT 837
QY 312 ValAlaSerSerSerLysArgThrThrIleSerPheGlnGluLeuAsnGlnLys-LysTy 331
Db 838 GTAGCATCAACAAATAGAAGAACACGATCCGGTTCGCAGAACTCAAAAAAAGGAGATA 897
QY 331 rArgThrProLeuLeuGlyIleGlyLeuLeuValLeuGlnAsnLeuSerGlyIleAs 351
Db 898 TTGGTTTCCATTGATGGTTGGAATTTGGGCTTCTTATGTTGCAACAACCTAGTGGCATTAA 957
QY 351 nGlyValLeuPheTyrAlaSerSerIlePheLysAlaAlaGlyValThrAsnSerAspLe 371
Db 958 TGGTGTATATTTCTATTCGCTACCATATTTCCGAAACTGCTGGAATTAATCAGGCAATAT 1017
QY 371 uAlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThrGlyValThrTrpLeu 390
Db 1018 AGCTACCTTTGGAGTTGGCTTTCATTTCAGGTCATGCTACTGCTTTTAAACCACATGGTTG 1075

RESULT 11
US-10-424-599-132426
; Sequence 132426, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 132426
; LENGTH: 2105
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(2105)
```

OTHER INFORMATION: unsure at all n locations
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_90588C.1
US-10-424-599-132426

Alignment Scores:
Pred. No.: 9.29e-104 Length: 2105
Score: 1139.50 Matches: 228
Percent Similarity: 70.55% Conservative: 69
Best Local Similarity: 54.16% Mismatches: 58
Query Match: 41.00% Indels: 67
DB: 16 Gaps: 2

US-10-051-909-36 (1-553) x US-10-424-599-132426 (1-2105)

QY 187 LyAspAlaSerPheLeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIle 206
Db 320 CAGGATTCCTCATTTTGTATATGGGAGCCTGCTGGAAGGTTTGGCGTCGGGATTAGC 379
QY 207 SerTyrThrValProValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeu 226
Db 380 TCATATTAGGTGCCTGTTTATATAGCTGAGATTGCACCTCAAACTTGAGAGGTGCCCTT 439
QY 227 GlySerValAsnGlnLeuSerValThrPheGlyIlePheLeuAlaTyrLeuGlyMet 246
Db 440 GGATCAGTGAACCAAGCTCTCTATTACTATTGGCATTATGCTGGCTTATCTTTGGGCTT 499
QY 247 PheIleProTyrArgLeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIlePro 266
Db 500 TTTGTCAACTGGAGAGTGCCTGCAATCTAGGAATTTGCCCTGTACAGTATTATACCT 559
QY 267 GlyLeuPhePheIleProGluSerProArgTyrLeu----- 278
Db 560 GGATTATTTTTCATACCTGAATCCCCAGATGGTTGGTATGGATATGTTATGCATTATAT 619
QY 278 ----- 278
Db 620 TGATCAAGTTTGTACTTTTCATCTTTGTTAAATACTGTTTCCTTTCAGAAAGGATG 679
QY 278 ----- 278
Db 680 GTAACCTTTTCTACTCTTGGTAATTATCTTTCTGCTTCTCTTTCATTTTATTTTATTTA 739
QY 279 -----AlaLysMetAsnLeuThrGluAspCy 287
Db 740 TTTTCTCTGGTTGGTGGGAATCGGGGAAGTGTAGGCTAAGATGGGATGACAGATGAGTT 799
QY 287 sGluThrSerLeuGlnValLeuArgGlyPheGluThrAspIleThrThrGluValAsnAs 307
Db 800 TGAGACTCTTTTGAAGTGTACGAGGATTTGACACAGATATATCTGTTGAAGTATATGA 859
QY 307 pIleLysArgAlaValAlaSerSerSerLysArgThrThrIleSerPheGlnGluLeuAs 327
Db 860 AATTAAGAGATCTGTGGCTTCAACGGGAAAGAGCTACAATCCGATTTGCAGATCTCAA 919
QY 327 nGlnLysLysTyrArgThrProLeuLeu-----LeuGlyIleGlyLe 341
Db 920 GAGGAAAGATATTGGTTCCCTTAATGGTATAATATGACACTGTAGTTGGTATTGGATT 979
QY 341 uLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSerIlePh 361
Db 980 ACTTGTCTTCAGCAGTTATCTGGTATCAATGGAGTTTGTCTTATTCAACTACCATCTT 1039
QY 361 eLysAlaAlaGlyValThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIleGlnVa 381
Db 1040 TGCAAAATGCAGGAATTTTCCTCCAGCGAAGCTGCTACTGTGGACTTGGTCCGTCAGGT 1099
QY 381 lLeuAlaThrGlyValThrThrTrpLeuLeuAspArgAlaGlyArgArgIleLeuLeuI 401
Db 1100 CATAGCAACTGGAATTTCCACGTGGTTGGTGGACAAAGTGGTCCGAGGCTGCTTCTAAT 1159
QY 401 eIleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSerValPhePheLeuLy 421

Db 1160 GATATCCTCATCTGTAAATGACTGTAGCCTTCTCATTTGTTTCAATAGCATTTTATCTGGA 1219
QY 421 sAspAsnIleSerGlnAspSerAsnSerTyrTyrIleLeuThrMetIleSerLeuValG 441
Db 1220 GGGGGTTGTATCTCAGGATTCACATCTATTAGCATGTTGGGAATAGTTTCTGTTGTTGG 1279
QY 441 yIleValSerPheValIleThrPheSerPheGlyMetGlyAlaIleProTyrLeuMetMe 461
Db 1280 ACTTGTGGTTATGGTCATTGGGTTCTCTCTGGTCTGGGACCATCCCTTGGCTTATAAT 1339
QY 461 tSerGluIleLeuProValSerIleLysSerLeuGlyGlySerIleAlaThrLeuAlaAs 481
Db 1340 GTCTGAGATACTTCCAGTGAATATAAAGGCCCTTGTCTGGCAGCATAGCGACAATGGGAA 1399
QY 481 nTrpLeuThrSerPheAlaIleThrMetThrThrAsnLeuMetLeuThrTrpSerValG 501
Db 1400 TTGGCTGATTTCGTGGGTGATCAGATGACCGCTAACCTTACTTTTGAATTGGAACAGTGG 1459
QY 501 yGlyThrPheLeuSerTyrMetValValSerAlaPheThrIleValPheValValLeuTr 521
Db 1460 AGGACATTTACAATCTACACAGTGGTAGCTGCCCTTACTATTGCTTTTATAGCATTTATG 1519
QY 521 pValProGluThrLysGly***AsnSerArgGlyAspThrIlePheValSerLeuSerIl 541
Db 1520 GGTTCCTGAGACCAAGGGAAG-AACATTGGAAGAAATTCAGTTTTCCTTCAGATAGATGT 1578
QY 541 e 541
Db 1579 T 1579
RESULT 12
US-10-425-114-5566
; Sequence 5566, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 5566
; LENGTH: 1027
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700464574_FLI
US-10-425-114-5566

Alignment Scores:
Pred. No.: 3.58e-104 Length: 1027
Score: 1139.00 Matches: 233
Percent Similarity: 99.57% Conservative: 0
Best Local Similarity: 99.57% Mismatches: 1
Query Match: 40.99% Indels: 1
DB: 16 Gaps: 0

US-10-051-909-36 (1-553) x US-10-425-114-5566 (1-1027)

QY 320 ThrIleSerPheGlnGluLeuAsnGlnLysLysTyrArgThrProLeuLeuGlyIle 339
Db 3 ACAATCAGTTTCAAGAATTAACCAAAAGAAATACCGCACGCCACTACTTCTAGGGATT 62
QY 340 GlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSer 359
Db 63 GGCCTACTTGTACTGCAAAATCTTAGTGAATCAACGGTGTACTGTTTATGCAAGTAGC 122

QY 360 IlePheLysAlaAlaGlyValThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIle 379
Db 123 ATCTTCAAAGCTGCAGGGGTTACAAACAGCGACTTGGCCACCTGTTCACTTGGTGTATT 182
QY 380 GlnValLeuAlaThrGlyValThrThrTrpLeuLeuAspArgAlaGlyArgAlaIleLeu 399
Db 183 CAGGTCCTTGCTACTGGAGTTACAACATGGCTGTAGACCGAGCTGGACGACGATCCTT 242
QY 400 LeuIleIleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSerValPhePhe 419
Db 243 CTCATTATTCTACCTCTGGCATGACTCTATGCCTTCTTGCCGTTTCTGTGATTATTTT 302
QY 420 LeuLysAspAsnIleSerGlnAspSerAsnSerTyrrIleLeuThrMetIleSerLeu 439
Db 303 CTCAGGATAACATTTCACAGGATTCTAATCATACTACTATCTTAACAATGATCTCCCTT 362
QY 440 ValGlyIleValSerPheValIleThrPheSerPheGlyMetGlyAlaIleProTrpLeu 459
Db 363 GTTGGTATTGTGCTTTGTTCATTACCTTCTCGTTTGGTATGGTGCCATTCATGGCTC 422
QY 460 MetMetSerGluIleLeuProValSerIleLysSerLeuGlyGlySerIleAlaThrLeu 479
Db 423 ATGATGCTGTAGATCTCCGGTTAGCATCAAGAGCCTTGGCGGAAGCATCGCAACACTG 482
QY 480 AlaAsnTrpLeuThrSerPheAlaIleThrMetThrThrAsnLeuMetLeuThrTrpSer 499
Db 483 GCCAACTGGCTGACATCTCTCGCCATAACAATGACGACGAACTTGTATGCTCAGTGGAGT 542
QY 500 ValGlyGlyThrPheLeuSerTyrrMetValValSerAlaPheThrIleValPheValVal 519
Db 543 GTTGGAGGCACCTTTCTCTCGTACATGGTTGTAGCGCCTTCCACATCGTTTTTGTGTC 602
QY 520 LeuTrpValProGluThrLysGly**AsnSerArgGlyAspThrIlePheValSerLeu 539
Db 603 CTTTGGTGCCGAGACGAGGAGGAG-AACTCTAGAGGAGATACAAATTTTCGTTTCGCTG 661
QY 540 SerIleGlnArgGlnLeuGlnTrpLeuProGluCysLeuSer 553
Db 662 AGCATTGAGCGTCAGCTGCAATGGTTGCCCGAGTGTTTATCT 703

RESULT 13

US-10-767-795-566
; Sequence 566, Application US/10767795
; Publication No. US20040181830A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Cao, Yongwei
; APPLICANT: Zhou, Yihua
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53534)B
; CURRENT APPLICATION NUMBER: US/10767,795
; CURRENT FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 117596
; SEQ ID NO 566
; LENGTH: 990
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: GOSHI-09MAY01-C14353_1

US-10-767-795-566
Alignment Scores:
Pred. No.: 6.61e-85 Length: 990
Score: 946.50 Matches: 196
Percent Similarity: 80.55% Conservative: 40
Best Local Similarity: 66.89% Mismatches: 53
Query Match: 34.06% Indels: 5
DB: 17 Gaps: 1
US-10-051-909-36 (1-553) x US-10-767-795-566 (1-990)

QY 53 GluSerGlySerAspHisAspGlyValLeuArgArgProLeuLeuAsnThrGlySerTrp 72

Db 124 GATGATGCTGAAGATGAAGAGGGGATCTAAAGGAAGCCATTCTACATACTGGAAGCTGG 183
QY 73 TyrArgMetSerSerArgGlnSerSerPheAlaProGlyThrSerSerMetAlaValLeu 92
Db 184 TATCGTATGGGTTCTAGGATGGGGTCTAGTATGTTAGGATCCCTCTCAA-----GCTCTT 237
QY 93 ArgGluSerHisValSerAlaPheLeuCysThrLeuIleValAlaLeuGlyProIleGln 112
Db 238 CGTGATAAATCCATCTCGGTTGTAGCTTGTGTTTGTATTGTTGCTTGGTCCCATCCAA 297
QY 113 PheGlyPheThrSerGlyPheSerSerProThrGlnAspAlaMetValArgAspLeuAsn 132
Db 298 TTTGGTTTCACTTCTGGTTACTCTTCTCCGACACAATCCGCAATCATCGAGGATCTTGA 357
QY 133 LeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyGlyMetValGly 152
Db 358 CTAACAGTCTCAGAGTTTCTCTATTGTTGTTCTTTATCAAATGTGGTGCCATGGTTGA 417
QY 153 AlaIleAlaSerGlyGlnMetAlaGluTyrrIleGlyArgLysGlySerLeuMetIleAla 172
Db 418 GCAATAGCCAGTGGTCAGATAGCCGAATGTATTGGACCGAAAGGGTCTTTAATGACTGCT 477
QY 173 AlaIle-ProAsnIleIleGlyTrpLeuAlaIleSerPheAlaLysAspAlaSerPheLe 192
Db 478 GCAATCCCTAATAATAATTGGATGGCTTGTATATCTTTTGCAAGAGATCTCTCTTTCT 537
QY 192 uTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTyrrThrValProVa 212
Db 538 TTACATGGGAAGGCTGTGCAAGGTTCTCGTCTCGGCATAATCTCTATACGGTGCCTGT 597
QY 212 lTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLe 232
Db 598 ATATATAGCTGAGATAGCACCTCAGAAATTTGAGGGGCGAGCTTGGGTTTCAGCGAATCAGCT 657
QY 232 uSerValThrPheGlyIlePheLeuAlaTyrrIleLeuLeuGlyMetPheIleProTrpArgLe 252
Db 658 ATCTGTCACTCTCGGAATAATGCTGGCCTATCTACTTCGACTTTTGTGAGTGCCGAT 717
QY 252 uLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPhePheIlePr 272
Db 718 ACTCGCCGTTTAGGAATACTTCTTGTGTACAATTTTGATACCTGCTCTATTTTTCATTCC 777
QY 272 oGluSerProArgTrpLeuAlaLysMetAsnLeuThrGluAspCysGluThrSerLeuGl 292
Db 778 AGAATCTCCTCGATGCTGGCAAACTGGGATGACAGAGATTTTCGAAACTTCTTTGCA 837
QY 292 nValLeuArgGlyPheGluThrAspIleThrThrGluValAsnAspIleLysArgAlaVa 312
Db 838 AGTTCTCAGGTGCTCCGATGCTGATATTTCTATTGAAGTGAATGAATCAACAGATCTGT 897
QY 312 lAlaSerSerSerLys-ArgThrThrIleSerPheGlnGluLeuAsnGlnLysLysTyra 332
Db 898 CTCATCAACCCTAGCAAGAACTACGATTCCGTTTG--CACATCTCAAGAAAGAGATATT 955
QY 332 rgThrProLeuLeuGlyIleGlyLeuLeuVal 343
Db 956 GGTCTCCATTGATGGTTGGAATTCGCTCATTATG 990

RESULT 14

US-09-938-842A-2331
; Sequence 2331, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIP1300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24

```
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2331
; LENGTH: 1473
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-2331

Alignment Scores:
Pred. No.:      4.81e-82      Length:      1473
Score:          920.50      Matches:      186
Percent Similarity: 61.69%      Conservative: 91
Best Local Similarity: 41.43%      Mismatches: 159
Query Match:    33.12%      Indels:      13
DB:              9          Gaps:         3

US-10-051-909-36 (1-553) x US-09-938-842A-2331 (1-1473)
QY  98 SerAlaPheLeuCysThrLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSer 117
Db  151 TCTGTTTCTCAGCACCTTGTGTCGGTATCAGGCTCTTCTGTACCGGTTGTGGCGTT 210
QY  118 GlyPheSerProThrGlnAspAlaMetValArgAspLeuAsnLeuSerIleSerGlu 137
Db  211 GGTTTTTCATCGGGTGCAAGCAGGGATTACCAAGATTATCTCTCTCCGTGCAGAA 270
QY  138 PheSerAlaPheGlySerLeuSerAsnValGlyGlyMetValGlyAlaIleAlaSerGly 157
Db  271 TACTCAATGTTCCGGTTCGATCTTGACATTAGGAGGCTTGATCGGTGCAGTATTCAGCGGT 330
QY  158 GlnMetAlaGluTyrIleGlyArgGlyGlySerLeuMetIleAlaAlaIleProAsnIle 177
Db  331 AAGTCGCTGATGTCTTGGGAAGAAACGGACGATGTGTTTTCGGAATTTCTTCTGTATC 390
QY  178 IleGlyTyrLeuAlaIleSerPheAlaLysAspAlaSerPheLeuTyrMetGlyArgLeu 197
Db  391 ACAGGCTGGCTTTGTGTAGCATGGCTCAGAAATGCAATGTGGCTGGACTGGGAAGATTG 450
QY  198 LeuGluGlyPheGlyValGlyIleIleSerTyrThrValProValTyrIleAlaGluIle 217
Db  451 TTACTTGGAAATCGGCGTTGGTATATTTAGCTACGTACGTCCGGTGTATATAGCCGAAATT 510
QY  218 SerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGly 237
Db  511 GCACCTAAACATGTCCGAGGATCGTTGTGTTCGCCAATCAGTTGATGCAAAATTCGCGGA 570
QY  238 IlePheLeuAlaTyrLeuLeuGlyMetPheIleProTyrArgLeuLeuAlaValIleGly 257
Db  571 ATTCACTCTTCTTCATCATTTGGCAATTTTATTCCATGGAGACTACTAACAGTAGTCGGA 630
QY  258 AlaLeuProCysThrMetLeuIleProGlyLeuPhePheIleProGluSerProArgTyr 277
Db  631 TTGGTGCCATGTGTGTTCCACGCTTTTGTATTATTTTTCATCCCGCAATCTCCCAAGATGG 690
QY  278 LeuAlaLysMetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPhe 297
Db  691 CTGGCGAAGTTAGGTCGTGATAAAGAATGCCGATCTTCGTGCAACGCCTTAGGGGATCT 750
QY  298 GluThrAspIleThrThrGluValAsnAspIleLysArgAlaValAlaSerSerLys 317
Db  751 GACGTCGATATTTCTCGTGAGCAACACAAATTAGATACCATTTGACATGACAGAAAC 810
QY  318 ArgThrThrIleSerPheGlnGluLeuAsnGlnLysLysTyrArgThrProLeuLeuLeu 337
Db  811 GGTGGTGAACATAAGATGTTGAAATTGTTTTCAGAGACCATACGCATATCCGTTAATTATC 870
QY  338 GlyIleGlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAla 357
```

```
Db  871 GGAGTTGGTTTAAATGTTTTTCACAACAATTGTGTGGGAGTCCGGTGTACCTATTATGCT 930
QY  358 SerSerIlePheLysAlaIaGlyValThrAsnSerAspLeuAlaThrCysSerLeuGly 377
Db  931 AGTAGCCTCTTCAACAAAGGAGGATTTCCA--AGTGCTATTGGCACATCCGTAAATAGCC 987
QY  378 AlaIleGlnValLeuAlaThrGlyValThrThrTrpLeuLeuAspArgAlaGlyArgArg 397
Db  988 ACAATTATGGTTCCAAAAGCAATGCTGGCAACAGTCCCTAGTCGATAAAATGGGGAGAGA 1047
QY  398 IleLeuLeuIleIleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSerValVal 417
Db  1048 ACGCTCCTAATGGCTTCTTGTGTTCTGCAATGGGTTTGAGTGTCTTCTTAAGTGT--- 1104
QY  418 PhePheLeuLysAspAsnIleSerGlnAspSerAsnSerTyrTyrIleLeu----- 434
Db  1105 -----TCTTACGGTTTCCAGTCGTTGGCATTTCTTCCAGAACTC 1143
QY  435 --ThrMetIleSerLeuValGlyIleValSerPheValIleThrPheSerPheGlyMet 453
Db  1144 ACTCCCATCTTCACCTGTCATCGCGCTTGTGGTGCACATTGTGTCAATTGCCATGGGAATG 1203
QY  454 GlyAlaIleProTyrLeuMetMetSerGluIleLeuProValSerIleLysSerLeuGly 473
Db  1204 GGAGGACTACCATGGATTATATGCTGAGATATTTCCGATGAATGTGAAAGTGTGAGCT 1263
QY  474 GlySerIleAlaThrLeuAlaAsnTrpLeuThrSerPheAlaIleThrMetThrThrAsn 493
Db  1264 GGGACCTTAGTTACTGTAAACCAATTGGTTATTTGGTTGGATTATCACATACACTTTCAAT 1323
QY  494 LeuMetLeuThrTrpSerValGlyGlyThrPheLeuSerTyrMetValValSerAlaPhe 513
Db  1324 TTTATGCTAGAAATGGAATGCATCAGGAATGTTCTCTCATCTTCTCAATGGTCTCCGCCAGT 1383
QY  514 ThrIleValPheValValLeuTrpValProGluThrLysGly***AsnSerArgGlyAsp 533
Db  1384 TCGATCGTATTTATATATACTTTTGTGTAACCTGAGACAAAAGCCGATCACTTGAAGAAATA 1443
QY  534 ThrIlePheValSerLeuSerIleGln 542
Db  1444 CAAGCACTGCTCAACAACCTCTGTGCAA 1470
```

RESULT 15

```
US-09-938-842A-2331
; Sequence 2331, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPT300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2331
; LENGTH: 1473
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-2331
```

```
Alignment Scores:      4.81e-82      Length:      1473
Pred. No.:              920.50      Matches:      186
Score:                  61.69%      Conservative: 91
Percent Similarity:
```

Best Local Similarity: 41.43% Mismatches: 159
Query Match: 33.12% Indels: 13
DB: 11 Gaps: 3

US-10-051-909-36 (1-553) x US-09-938-842A-2331 (1-1473)

```
QY 98 SerAlaPheLeuCysThrLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSer 117
Db 151 TCTGTTTCTCCAGCACCTTTGTTGCCGTATCAGGCTCCTTCTGTACCGGTTGTGGCGTT 210
QY 118 GlyPheSerSerProThrGlnAspAlaMetValArgAspLeuAsnLeuSerIleSerGlu 137
Db 211 GGTTTTTCATCGGGTGCACAAAGCAGGATACCAAGATTATCTCTCTCCGTTGCAGAA 270
QY 138 PheSerAlaPheGlySerLeuSerAsnValGlyGlyMetValGlyValAlaIleAlaSerGly 157
Db 271 TACTCAATGTTCCGGTCGATCTTGACATTAGGAGGCTTGATCGGTGCAGTATTCAGCGGT 330
QY 158 GlnMetAlaGluTyrIleGlyArgLysGlySerLeuMetIleAlaAlaIleProAsnIle 177
Db 331 AAAGTCGCTGATGTTGGGAAGAAAACGACGATGTTGTTTTGCCAATCTTCTGTATC 390
QY 178 IleGlyTrpLeuAlaIleSerPheAlaLysAspAlaSerPheLeuTyrMetGlyArgLeu 197
Db 391 ACAGGCTGGCTTTGTGTAGCATTTGGCTCAGAATGCAATGTGGCTGGACTGTGGAAGATTG 450
QY 198 LeuGluGlyPheGlyValGlyIleIleSerTyrThrValProValTyrIleAlaGluIle 217
Db 451 TTACTTGGAAATCGGCGTTGGTATATTTAGCTAGCTGATTCGGGTGATATAGCCGAAAT 510
QY 218 SerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGly 237
Db 511 GCACCTAAACATGTCGAGGATCGTTTGTGTGCCAATCAGTTGATGCAAAATTTGCCGA 570
QY 238 IlePheLeuAlaTyrLeuLeuGlyMetPheIleProTrpArgLeuLeuAlaValIleGly 257
Db 571 ATTTCACCTCTTCTTCATCATTTGGCAATTTTATTCATGGAGACTACTAACAGTAGTCGGA 630
QY 258 AlaLeuProCysThrMetLeuIleProGlyLeuPhePheIleProGluSerProArgTrp 277
Db 631 TTGGTGCCATGTGTGTTCCACGTCTTTTGTATTATTTTCATCCCGCAATCTCCAAGATGG 690
QY 278 LeuAlaLysMetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPhe 297
Db 691 CTGGCGAAGTTAGTGCTGATATAAAGAAAGCGGATCTTCGTTGCAACGCTTAGGGGATCT 750
QY 298 GluThrAspIleThrGluValAsnAspIleLysArgAlaValAlaSerSerLys 317
Db 751 GACGTCGATATTTCTCGTGAAGCAAAACAACAAATTAGAGATACCAATTGACATGACAGAAAAC 810
QY 318 ArgThrThrIleSerPheGlnGluLeuAsnGlnLysLysTyrArgThrProLeuLeu 337
Db 811 GGTGGTGAAACTAAGATGCTCTGAATTTGTTTCAGAGACGATACGCAATATCCGTTAATTATC 870
QY 338 GlyIleGlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAla 357
Db 871 GGAGTTGGTTTAAATGTTTGTCAACAATTTGTGGAGCTCCGGTGTACCTATTATGCT 930
QY 358 SerSerIlePheLysAlaAlaGlyValThrAsnSerAspLeuAlaThrCysSerLeuGly 377
Db 931 AGTAGCCTCTTCAACAAGAGGAGGATTTCCCA--AGTGCTATTGGCACATCCGTAATAGCC 987
QY 378 AlaIleGlnValLeuAlaThrGlyValThrTrpLeuLeuAspArgAlaGlyArgArg 397
Db 988 ACAATTATGGTTCCAAAAGCAATGCTGGCAACAGCTCCTAGTCGATAAAATGGGGAGGAGA 1047
QY 398 IleLeuLeuIleIleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSerValVal 417
Db 1048 AGCTCCTAATGGCTTCTTGTCTGCAATGGGTTTGAGTGTCTTGTCTCTTAAGTGT--- 1104
QY 418 PhePheLeuLysAspAsnIleSerGlnAspSerAsnSerTyrTyrIleLeu----- 434
Db 1105 -----TCTTACGGTTTCCAGTCGTTTGGCAATCTTCCAGAACTC 1143
```

```
QY 435 ---ThrMetIleSerLeuValGlyIleValSerPheValIleThrPheSerPheGlyMet 453
Db 1144 ACTCCCATCTTCACTTGCATCGCGTCTTGGGTACACATTGTGTCAATTTGCCATGGGAATG 1203
QY 454 GlyAlaIleProTrpLeuMetMetSerGluIleLeuProValSerIleLysSerLeuGly 473
Db 1204 GGAGGACTACCATGGATTATAATGGCTGAGATATTTCCGATGAATGTGAAAGTGTACGCT 1263
QY 474 GlySerIleAlaThrLeuAlaAsnTrpLeuThrSerPheAlaIleThrMetThrThrAsn 493
Db 1264 GGGACCTTAGTTACTGTAAACCAATTGGTTATTGGTTGGATTATCACATACACTTTTCAAT 1323
QY 494 LeuMetLeuThrTrpSerValGlyGlyThrPheLeuSerTyrMetValValSerAlaPhe 513
Db 1324 TTTATGCTAGATGGAATGCATCAGGAATGTTCTCTCATCTTCTCAATGGTCTCCGCCAGT 1383
QY 514 ThrIleValPheValValLeuTrpValProGluThrLysGly***AsnSerArgGlyAsp 533
Db 1384 TCGATCGTATTTATATATACTTTTGGTACCTGAGACAAAGGCCGATCACTTGAAGAAATA 1443
QY 534 ThrIlePheValSerLeuSerIleGln 542
Db 1444 CAAGCACTGCTCAACAACACTCTGTGCAA 1470
```

Search completed: October 14, 2004, 00:03:56
Job time : 825.832 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2004, 11:33:22 ; Search time 14.714 Seconds
(without alignments)
3616.147 Million cell updates/sec

Title: US-10-051-909-36
Perfect score: 2779
Sequence: 1 PSSSSSFRPAGKKKKKNQG.....TIFVSLSIQRLQLWLPCLSL 553

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1757.5	63.2	490	T14545	probable sugar tra
2	1734.5	62.4	487	E96782	hypothetical prote
3	920.5	33.1	496	T52132	probable sugar tra
4	824	29.7	483	D96589	hypothetical prote
5	820.5	29.5	474	B86221	hypothetical prote
6	808	29.1	454	A86221	hypothetical prote
7	748.5	26.9	461	G85059	probable sugar tra
8	729	26.2	457	H85059	probable sugar tra
9	671.5	24.2	378	G86220	hypothetical prote
10	626	22.5	490	F86220	hypothetical prote
11	621	22.3	348	D84922	probable sugar tra
12	578.5	20.8	521	G84864	probable membrane
13	575.5	20.7	524	A31318	glucose transporte
14	566.5	20.4	560	T51485	sugar transporter-
15	563.5	20.3	496	A41264	glucose transport
16	560.5	20.2	493	A41751	glucose transport
17	557.5	20.1	457	E70070	metabolite transpo
18	556.5	20.0	511	A84537	probable sugar tra
19	553	19.9	509	A32101	glucose transport
20	551.5	19.8	511	H84536	probable sugar tra
21	550	19.8	509	A33801	muscle-fat glucose
22	549	19.8	523	S06920	glucose transport
23	548.5	19.7	493	S38981	glucose transport
24	542.5	19.5	496	A31986	glucose transporte
25	542.5	19.5	522	A31556	glucose transport
26	540	19.4	482	B69803	metabolite transpo
27	538	19.4	580	D86426	hypothetical prote
28	536.5	19.3	508	G84564	probable sugar tra
29	531.5	19.1	508	B30310	glucose transport

30	531	19.1	473	2	G69789	sugar transporter
31	525	18.9	493	2	A85433	sugar transporter
32	523	18.8	464	2	AC0877	galactose-proton s
33	522	18.8	464	2	F65079	galactose-proton s
34	522	18.8	549	2	T14606	probable sugar tra
35	517.5	18.6	492	2	I45902	glucose transporte
36	517	18.6	461	2	D70073	metabolite transpo
37	515.5	18.5	464	2	C91106	galactose-proton s
38	515.5	18.5	464	2	F85951	galactose-proton s
39	515.5	18.5	513	2	T01506	probable hexose tr
40	512.5	18.4	492	2	A27217	glucose transport
41	512.5	18.4	492	2	A25949	glucose transport
42	511.5	18.4	492	2	S09705	glucose transport
43	511.5	18.4	523	2	S25015	monosaccharide tra
44	510.5	18.4	492	2	A30797	glucose transport
45	508	18.3	469	2	H97064	probable sugar-pro

ALIGNMENTS

RESULT 1

T14545
probable sugar transporter protein - beet
C;Species: Beta vulgaris (beet)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C;Accession: T14545
R;Chiou, T.J.; Bush, D.R.
Plant Physiol. 110, 511-520, 1996
A;Title: Molecular cloning, immunochemical localization to the vacuole, and expression
A;Reference number: Z18131; MUID:96351183; PMID:8742332
A;Accession: T14545
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-490 <CHI>
A;Cross-references: UNIPROT:Q39416; EMBL:U43629; NID:G1209755; PIDN:AAB53155.1; PID:G12
A;Experimental source: tonoplast
C;Superfamily: glucose transport protein
C;Keywords: transmembrane protein

Query Match 63.2%; Score 1757.5; DB 2; Length 490;
Best Local Similarity 66.2%; Pred. No. 9.5e-118;
Matches 337; Conservative 76; Mismatches 73; Indels 23; Gaps 3;

QY	39	MGGGNNRGAGAGEESGDHGVLRPLNTGTSWYRMSSRQSSFAPGTSSMAVLRESHVS	98
DB	1	MSSDSEAGLGGG-----GDLRKPFLHTGTWYRMGSRQSSL---MGSSQVIRESSIS	49
QY	99	AFLCTLIVALGPIQGFSTGFSPTQDAMVRDLNLSISEFSAFGSLSNVGMVGAIASGQ	158
DB	50	VLACVLIVALGPIQGFSTAGYSSPTQSAITNELGLSVAEYSWFGSLSNVGMVGAIASGQ	109
QY	159	MAEYIGRKSLMIAAIPNIIIGWLAIISFAKADASFLYMGRLLEGFGVGIISYTPVYIAEIS	218
DB	110	ISEYIGRKSLMIAAIPNIIIGWLAIISFAKADSSFLYMGRLLEGFGVGIISYTPVYIIEIA	169
QY	219	PQNMKGALGSVNQLSVTFGIFLAYLLGMFIPWRLLAIVIGALPCTMLIPGLFFIPESPRWL	278
DB	170	PQNLKGALGSVNQLSVTIGIMLSYMLGLFVPWRILAVLGILPCTILIPGLFFIPESPRWL	229
QY	279	AKMNLTEDCETSLQVLRGFETDITTEVNDIKRAVASSSKRTTISFQELNQKRYTPLLLG	338
DB	230	AKMGMEEFVSLQVLRGFDTDISLEVNEIKRSVASSSKRTTIRFAELRQRRYWLPLMIG	289
QY	339	IGLLVLQNLSGINGVLFYASSIFKAAGVTNSDLATCSLGAIQVATGVTTWLLDRAGRI	398
DB	290	NGLLILQLSGINGVLFYSSSTIFKEAGVTSSNAATFGLGAVQVATVTTWLVDKSGRRL	349
QY	399	LLIISTSGMTLCLLAVSVVFFLKDNIQSQDSNSYYILTMISLVGIVSFVITFSFGMGAIPW	458
DB	350	LLIVSSSGMTLSLLVAMSFLLKEMVDSDETWYSVFSILSVGVVAMVVTFSLGIGAIPW	409
QY	459	LMSEILPVSIKSLGGSIAATLANWLTSPAITMTTNLMTWSVGTFSLSYMVVSFTIVFV	518

Db 410 IIMSEILPINIKGLAGSIATLANFWAVIVTMTANIMLSWNSGGTFSIYMVVCFTVAFV 469
QY 519 VLVVPETKGXNSRGDTIFVLSIQRLQW 547
Db 470 VIWVPETKGR-----LEEIQW 486

RESULT 2
E96782
hypothetical protein F22H5.6 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: E96782
R;Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: E96782
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-487 <STO>
A;Cross-references: UNIPROT:Q9PRL3; GB:AE005173; NID:g10092276; PIDN:AAG12689.1; GSPDB:G
C;Superfamily: glucose transport protein

Query Match 62.4%; Score 1734.5; DB 2; Length 487;
Best Local Similarity 70.2%; Pred. No. 4.1e-116;
Matches 334; Conservative 67; Mismatches 68; Indels 7; Gaps 2;

QY 52 BEGSDHGDVLRRLPLNTGTSWYRMSRQSSFAFCTSSMAVLRSHVSAPLCTLIIVALGPI 111
Db 8 BEARND---LRRPFHTGTSWYRMSRQSSM---MGSSQVIRDSSISVLACVLIVALGPI 60

QY 112 QFGFTSGFSSPTQDAMVRDLNLSISEFSAFGSLNVGVMGAIASGQMAEYIGRKGSLMI 171
Db 61 QFGFTCGYSSPTQAAITKDLGLTVSEYSVFGSLNVGAMVGAIASGQIAEYIGRKGSLMI 120

QY 172 AAIPNIIGWLAIISFAKDASFLYMGRLLEGFGVGIISYTPVYIAEISIPQNMRGALGSVQ 231
Db 121 AAIPNIIGWLICISFAKDTSPLYMGRLLEGFGVGIISYTPVYIAEIAIPQNMRGGLSVQ 180

QY 232 LSVTFGIFLAYLLGMFIPWRLLAVIGALPCTMLIPGLFFIPESPRWLAKMNLTEDCETSL 291
Db 181 LSVTIGIMLAYLLGLFVPRILAVLILPCTLLIPGLFFIPESPRWLAKMGWTFETSL 240

QY 292 QVLRGFETDITTEVNDIKRAVASSSKRTTISFQELNQKRYRTPLLLGILLVLQNLGIN 351
Db 241 QVLRGFETDITVEVNEIKRSVASSTKRTNTRVFDLKRRIYFFPLMVGILLVLQQLGIN 300

QY 352 GVLFYASSIFKAAAGVTNSDLATCSLGAIOVLATGVTTWLLDRAGRILLIISTSGMTLCL 411
Db 301 GVLFYSSSTIFESAGVTSSNAATFGVGAIOVVAISTAISTWLVDKAGRRLLLTSSVGMTISL 360

QY 412 LAVSVVFFLKDNIQSNSYIILTMISLVGIVSFVITPSFGMGAIPWLMMSAILPVSIKS 471
Db 361 VIVAAAFYLKEFVSPSDMYSLWSILSVGVVAMVVFFSLGMGPILPWLIMSEILPVNIKG 420

QY 472 LGGSIATLANWLTSPAITMTNMLTWSVGGTFLSYMVVSAPTIVFVVLWVPETKG 527
Db 421 LAGSIATLANWFFSWLITMTANLLAWSSGGTFTLYGLVCAFTTVFVTLWVPETKG 476

RESULT 3

T52132
probable sugar transporter protein ERD6 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 09-Jul-2004
C;Accession: T52132
R;Kiyosue, T.; Abe, H.; Yamaguchi-Shinozaki, K.; Shinozaki, K.
Biochim. Biophys. Acta 1370, 187-191, 1998
A;Title: ERD6, a cDNA clone for an early dehydration-induced gene of Arabidopsis, encode
A;Reference number: Z25973; MUID:98213606; PMID:9545564
A;Accession: T52132
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-496 <KIY>
A;Cross-references: UNIPROT:O65799; EMBL:D89051; PIDN:BAA25989.1
C;Genetics:
A;Gene: ERD6
C;Superfamily: glucose transport protein
C;Keywords: transmembrane protein; transport protein

Query Match 33.1%; Score 920.5; DB 2; Length 496;
Best Local Similarity 41.4%; Pred. No. 4.7e-58;
Matches 186; Conservative 91; Mismatches 159; Indels 13; Gaps 3;

QY 98 SAFLCTLIIVALGPIQFGFTSGFSSPTQDAMVRDLNLSISEFSAFGSLNVGGMVGAIASG 157
Db 57 SVFLSTFVAVSGSFCCTGCGVGFSSGAQAGITKDLSLVAEYSMFGSILTLGLIGAVFSG 116

QY 158 QMAEYIGRKGSLMIAAIPNIIGWLAIISFAKDASFLYMGRLLEGFGVGIISYTPVYIAEI 217
Db 117 KVADVLRKRTMLFCFFCITGWLCLVALAQNAMWLDGRLLLGIGVGFYVIPVYIAEI 176

QY 218 SPQNMRGALGSVNQLSVTFGIFLAYLLGMFIPWRLLAVIGALPCTMLIPGLFFIPESPRW 277
Db 177 APKHVRGSEVFANQLMQNCGISLFFIIGNFIPWRLLTVGLVPCVHFVFCFLFFIPESPRW 236

QY 278 LAKMNLTEDCETSLQVLRGFETDITTEVNDIKRAVASSSKRTTISFQELNQKRYRTPLLL 337
Db 237 LAKLRDKECRSSQLRGRSDVDISREANTIRDITDMTENGGETKMSLFFQRYAYPLII 296

QY 338 GIGLLVLQNLSGINGVLFFYASSIFKAAAGVTNSDLATCSLGAIOVLATGVTTWLLDRAGR 397
Db 297 GVGLMFLQQLCGSSGVTYASSLNFKNKGFP-SAIGTSVIATIMVPKAMLATVLVDKMGRR 355

QY 398 ILLIISTSGMTLCLLAVSVVFFFLKDNISQDSNSYIL----TMISLVGIVSFVITPSFGM 453
Db 356 TLLMASCSAMGLSALLSV-----SYGFSFGILPELTPIFTCIGVLGHIVSFAMGM 407

QY 454 GAIPWLMMSAILPVSIKSLGSGSIATLANWLTSPAITMTNMLTWSVGGTFLSYMVVSAP 513
Db 408 GGLPWIIMAEIIFPMNVKVSAGTLVTVTNWLFGWIITYTFNFMLEWNASGMFLIFSVMVAS 467

QY 514 TIVFVVLWVPETKGXNSRGDTIFVLSLIQ 542
Db 468 SIVFIYFLVPETKGRSLEIQALLNNSVQ 496

RESULT 4
D96589
hypothetical protein T22H2.15 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: D96589
R;Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712

Db 81 MGAIFSGKADLMGRKGMWFAQIFCFGWAVAKADSMWLDIGRLSTGFAVGLLSYV 140
QY 210 VPVYIAEISPQNMRGALGSVNQLSVTFGIFLAYLLGMFIPWRLLAIVIGALPCTMLIPGLF 269
Db 141 IPVYIAEITPKHVRGAFVFAVQALMQSCGLSLFYVIGNFVHRNLALIGLIPCALQVVTLF 200
QY 270 FIPESPRWLAKMNLTEDETSLQVLRGFETDITTEVNDIKRAVASSSKRTTISFQELNQK 329
Db 201 FIPESPRLLGKWGHEKCRASLQSLRGDDADISEANTIKETMILFDEGPKSRVMDLQF 260
QY 330 KYRTPLLLGIGLLVLQNLGNGVLFYASSIFKAAGVTNSDLATCSLGAIQVLATGVTTW 389
Db 261 RYAPSVVIGVGLMLLQQLSGSSGLMYVGVDFDKGFPSS-IGSMILAVIMIPKALLGLI 319
QY 390 LLDRAGRRIILLII-----STSGMTLCLLAVSVVFFLKDNIQSNSYYIL----TWIS 438
Db 320 LVEKMGRRPLLLMNDLYLQASTGCMCFSSLSSFSFCFR-----SYGMLDELTPIFT 371
QY 439 LVGIVSFVITFSFGMGAIPWLMSEILPVSIKSLGGSIAATLANWLTSAITMTTNMLTW 498
Db 372 CIGVVGFISSFAVGMGLPWIIIMSEIFPMNVKVSAGTLVLANWSFGWIVAFAYNFMLEW 431
QY 499 SVGGTFLSYMVV 510
Db 432 NASGINTYYIII 443

RESULT 7
G85059
probable sugar transporter [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: G85059
R;anonymouse, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring Nature 402, 769-777, 1999
A;Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A;Reference number: A85001; MUID:20083488; PMID:10617198
A;Accession: G85059
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-461 <STO>
A;Cross-references: UNIPROT:Q9M100; GB:NC_001268; NID:g7267233; PIDN:CAB80840.1; GSPDB:G85059

Query Match 26.9%; Score 748.5; DB 2; Length 461;
Best Local Similarity 35.5%; Pred. No. 7.8e-46;
Matches 168; Conservative 91; Mismatches 171; Indels 43; Gaps 6;

QY 81 SFAPGTSSMAVLRSHVSA---FLCTLIVALGPQIFGFTSGFSSPTQDAMVRDLNLSISE 137
Db 15 SASPNKSSLLSEISNASTRPFVLAFTVSGCGALSFGCI----- 53

QY 138 FSFAGSLSNVGMVGAIASQMAEYIGRKGSMLIAAIPNIIGWLAISFAKDASFLYMGRL 197
Db 54 FSFFGSILTVGLILGALICGLADLVGRVYTIWITNILVLIGWLAIAFAKDVRLDLGRL 113

QY 198 LEGFGVGIISYTPVYIAEISPQNMRGALGSVNQLSVTFGIFLAYLLGMFIPWRLLAIVIG 257
Db 114 LQGISVGISSYGLPIYISELAPNLRGAASSLMQLFVGVLGSAFYALGTAVAWRSAILG 173

QY 258 ALPCTMLIPGLFFIPESPRWLAKMNLTEDETSLQVLRGFETDITTEVNDIKRAVASSSK 317
Db 174 SIPSVLVPLPLFFIPESPRWLAKVGEKEVEGVLLSLRGAKSDVSDEAATILEYTKHVEQ 233

QY 318 RTTIS--FQELNQKKYRTPLLLGIGLLVLQNLGNGVLFYASSIFKAAGVTNSDLATCS 375
Db 234 QDIDSRGFFKLQKRYALPLTIGVVLISMPQLGGLNGYFTDTFTSTGV-SSDIGFIL 292

QY 376 LGAIQVLATGVTTWLLDRAGRRILLIISTSGMTLCLLAVSVVFFLKDNIQSNSYYILT 435

Db 293 TSIVQMTGGVLGVLLVDISGRRSLLLSQAGNFGCLATAISFFLQKXNCWETGT---P 348
QY 436 MISLVGIVSFVITFSFGMGAIPWLMSEILPVSIKSLGGSIAATLANWLTSAITMTTNLM 495
Db 349 IMALISVMVYFGSYGLGMGPIPIWIIASEIYPVDVKAAGTVCNLTSSISWLVYTSFNFL 408
QY 496 LTWSVGGTFLSYMVVSAFTIVFVVLWVPETKG-----XNSRGDTIF 536
Db 409 LQWSSTGTFMFMFATVMGLGFVFTAKLVPETKSKSLEEIQSAFTDSTSEDSTIF 461

RESULT 8
H85059
probable sugar transporter [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: H85059
R;anonymouse, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring Nature 402, 769-777, 1999
A;Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A;Reference number: A85001; MUID:20083488; PMID:10617198
A;Accession: H85059
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-457 <STO>
A;Cross-references: UNIPROT:Q9M029; GB:NC_001268; NID:g7267234; PIDN:CAB80841.1; GSPDB:G85059

Query Match 26.2%; Score 729; DB 2; Length 457;
Best Local Similarity 35.6%; Pred. No. 1.9e-44;
Matches 161; Conservative 91; Mismatches 170; Indels 30; Gaps 5;

QY 80 SSFAPGTSSMAVLRSHVSAFLCTLIV-ALGPQIFGFTSGFSSPTQDAMVRDLNLSISEF 138
Db 11 SSTSSSSSLLSEISNACTRPFVLAFTVSGCGAFAGCI-----Y 49

QY 139 SAFGSLSNVGMVGAIASQMAEYIGRKGSMLIAAIPNIIGWLAISFAKDASFLYMGRL 198
Db 50 SLFGSILTVGLILGALICGLTDLVGRVKTWITNILFVIGWLAIAFAKGVWLLDLGRL 109

QY 199 EGFGVGIISYTPVYIAEISPQNMRGALGSVNQLSVTFGIFLAYLLGMFIPWRLLAIVIG 258
Db 110 QGISIGISVYLGVPVYITEIAPNLRGAASSFAQLFAGVGISVYALGTIVAWRNLAIGC 169

QY 259 LPCTMLIPGLFFIPESPRWLAKMNLTEDETSLQVLRGFETDITTEVNDIKRAVASSSKR 318
Db 170 IPSLMVPLPLFFIPESPRWLAKVGEKEVEAVLSSLRGKSDVSDEAAEILEYTEHVKKQ 229

QY 319 TTI---SFQELNQKKYRTPLLLGIGLLVLQNLGNGVLFYASSIFKAAGVTNSDLATCS 375
Db 230 QDIDDRGFFKLQKRYAFSLTIGVVLIALPOLGGLNGYSFYTDSIFISTGV-SSDFGFIS 288

QY 376 LGAIQVLATGVTTWLLDRAGRRILLIISTSGMTLCLLAVSVVFFLKDNIQSNSYYILT 435
Db 289 TSVQMFQGGILGTLVDVSGRRRTLLVVSQAGMFLGCLTTAISFFLKENHCWETGT---P 344

QY 436 MISLVGIVSFVITFSFGMGAIPWLMSEILPVSIKSLGGSIAATLANWLTSAITMTTNLM 495
Db 345 VLALFSVMVYFGSYGSGMGSIPWIIASEIYPVDVKAAGTMCNLTSSISAWLVAYSFSYL 404

QY 496 LTWSVGGTFLSYMVVSAFTIVFVVLWVPETKG 527
Db 405 LQWSSTGTFMFMFATVAGLGFVFIKLVLPETKG 436

RESULT 9
G86220
hypothetical protein [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: G86220
R;Theologias, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marzialli,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PIDN:AAB70414.1; GSPDB:GN
A;Accession: G86220
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-378 <STO>
A;Cross-references: UNIPROT:O04040; GB:AE005172; NID:g2342688; PIDN:AAB70414.1; GSPDB:GN
C;Genetics:
A;Map position: 1

Query Match 24.2%; Score 671.5; DB 2; Length 378;
Best Local Similarity 34.3%; Pred. No. 1.9e-40;
Matches 148; Conservative 66; Mismatches 125; Indels 93; Gaps 4;
QY 119 FSSPTQDAMVRDLNLSISEFSAFGSLSNVCGMVGMAIASQMAEYIGRK----- 166
Db 3 YSSPAQSKIMEELGSLVADYSFPTSVMTLGGMTITAVFSGKISALVGRQVSIYYCWTGSG 62
QY 167 -----GSLMIAAIPNIIGWLAISPAKDAFLYMGRLLEGVGGIISYTVPVYIA 215
Db 63 LSPFEFFWIFYDDTMWISDVCCIFGWLAVAFADHIIIMLTGRLFLGFGVGLISYVVPYIA 122
QY 216 EISPNMRGALGSVNQLSVTFGIFLAYLLGMFIPWRLLAVIGALPCTMLIPGLFPIESP 275
Db 123 EITPKTFRGGESYNSQLLQCLGISLMPFTGNFFHWRTLALISAIPSAFQVICLFFIPESP 182
QY 276 RWLAKMNLTEDETSLOVLRGFETDITTEVNDIKRAVASSSKRTTISFQELNQKRYTPL 335
Db 183 RWLAMYGDQDELEVSLLKLRGENSDILKEAAEIR----- 216
QY 336 LLGIGLLVLQNSGIVGVLFYASSIFKAAAGVTNSDLATCSLGAIOVLATGVTWLLDRAG 395
Db 217 -IGLGLMLLQFCGSAAISAYARIFDKAGFP-SDIGTTILAVILIPQSIIVMLTVDRWG 274
QY 396 RRLIIISTGSMTCCLLAVSVFVLKDNISQDSNSYYILTMISLVGIVSFVITFSFGMGA 455
Db 275 RRLPL-----MGVVSFSGIGLGG 292
QY 456 IPWLMMSLIPVSIKSLGSGSIATLANWLTSAITMTNMLTWSVGGTFLSYMVVSAFTI 515
Db 293 LPWVIMSEIFPNVKITAGSLVTMSNPFNFNIIYISFNFMIQMSASGTTFIFSGVSLVTI 352
QY 516 VFVVLWVPETKG 527
Db 353 VFIWTLVPETKG 364

RESULT 10
F86220
hypothetical protein [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: F86220
R;Theologias, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marzialli,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PIDN:AAB70415.1; GSPDB:GN
A;Accession: F86220
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-490 <STO>
A;Cross-references: UNIPROT:O04041; GB:AE005172; NID:g2342689; PIDN:AAB70415.1; GSPDB:GN
C;Genetics:
A;Map position: 1

Query Match 22.5%; Score 626; DB 2; Length 490;
Best Local Similarity 30.6%; Pred. No. 4.6e-37;
Matches 165; Conservative 84; Mismatches 162; Indels 128; Gaps 12;
QY 53 EGSDDHGVLRRLNLTGWSYRMSSRQSSFAFGTSSMAVLRSHVSAFLCT----- 103
Db 2 ESGS-----MKTPLVNNQEEARSSSITCGLLLSVAV-TGSFVYGACSSWSTNMPQ 55
QY 104 --LIVALGPIQFGFTS-----GFSPTQDAMVRDLNLSISEFSAFGSLN 146
Db 56 IGIFLIKICFEFVDVNLWPNFVHYHNLKMSYSSPAQSKIMEELGSLVADYSFETSVMT 115
QY 147 VGMVGAIASQMAEYIGRKSLMIAAIPNIIGWLAISFA-----KDAS 190
Db 116 LGGMITAFAFGKIAAIVGRQTMWADVFCIFGWLAVAFAHYFLIKLTFHLLWFIKDKM 175
QY 191 FLYMGRLLGFGVGIISYTVPVYIAEISPNMRGALGSVNQLSVTFGIFLAYLLGMFIPW 250
Db 176 LLENIGRGFLGFGVGLISYVVPYIAEITPKAFRGGSFNSQLLQSPGISLMPFTGNFFHW 235
QY 251 RLLAVIGALPCTMLIPGLFPIESPRL-----AKMNLTEDETSLOVLRGFETDITTE 304
Db 236 RTLALLSAIPCGIQMICLFFIPESPRLVIHTLKAMYGRELEVTILKRLRGENDILEE 295
QY 305 VNDIKRAVASSSKRTTISFQELNQKRYTPLLIGLGLLVQLNLSGIVGVLFYASSIFKAA 364
Db 296 AAER-----IGLGLMLLQFCGSSAISAYARIFDTA 328
QY 365 GVTVNSDLATCSLGAIOVLATGVTWLLDRAGRIIL-----IISTGSMTCCLLAVSVVFFLK 421
Db 329 GTA---IYCHVCRPPMTFTTDELFYRSVHMFIFNPWLLSSG---LVILTILFF-- 379
QY 422 DNISQDSNSYYILTMISLVGIVSFVITFSFGMGAIPWLMMSLIPVSIKSLGSGSIATLAN 481
Db 380 -----GYVLSFGIGLGLPVMIMSEVFPVNVKITAGSLVTVS 417
QY 482 WLTSFAITMTNMLTWSV-----GGTFLSYMVVSAFTIVFVVLWVPETKG 527
Db 418 WFFSWIIIFSNFNMQWSAFEYITNTKMLYFDRTYFIFAGVSLMSFVFWTLVPETKG 476

RESULT 11
D84922
probable sugar transporter [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: D84922
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PIDN:AAD13706.1; GSPDB:GN
A;Accession: D84922
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-348 <STO>
A;Cross-references: UNIPROT:Q9ZU87; GB:AE002093; NID:g4249409; PIDN:AAD13706.1; GSPDB:GN
C;Genetics:
A;Gene: At2g48020
A;Map position: 2

Query Match 22.3%; Score 621; DB 2; Length 348;

Best Local Similarity 33.4%; Pred. No. 6.9e-37;
Matches 143; Conservative 70; Mismatches 95; Indels 120; Gaps 8;

QY	100	FLCTLI	VALGPIQFGFTSGFSSPTQDAMVRDLNLSISEFSAPGSLSNVGMVGATASGQM	159
Dd	29	YLSTFVAVCGSFAFGSCAGYSPPAQAAIRNDLSLTIAEFSLFGSLLTFGAMIGAITSGPI	88	
QY	160	AEYIGRKGSIMTAAIPNIIGWLAIISPAKDAFLYMGRLLLEGFGVGIISYTPVVIABISP	219	
Dd	89	ADLVGRKGGVVA-----LDLGRLATGYCMGAFSYVVPIFIABEIAP	128	
QY	220	QNMRGALGSVNQLSVTFGIFLAYLLGMFIPWRLLAVIGALPCTMLIPGLFFIPESPWLA	279	
Dd	129	KTRFGALTTLNQ-----A	141	
QY	280	KMNLTEDCETSLQVLRGFETDITTEVNDIKRAVASSSKRTTISFQELNQKYRTPILLGI	339	
Dd	142	KVGRDTEFEAALRKLGRKKADISEBAEIQ-----IAF	174	
QY	340	GLLVQLNLSGINGVLFYASSIFKAAGVTNSDLATCSLGAIOVLATGVTTWLLDRAGRIL	399	
Dd	175	GLMVFOQGGINGICFYTSSIFEQAQFP-TRLGMIIIVAVLQVSATG-----	219	
QY	400	LIIISTGMTLCLLAVSVVFFLKDNISQDSNSYYILMTISLVGIVSFVITFSGFMGAIPWL	459	
Dd	220	LVIG-----CLIA-AVSFYCLKVH---DMAHEAVPVULAUVVGIMVMYIGSFSAGMGAMPVW	268	
QY	460	MMSEILPVSIKSLGGSIA TLANWLTSFAITMTTNLMLTWSVGGTFLSYMVVSAFTIVFW	519	
Dd	269	VMSEIIPPINIKGVAGGMATLVNWF GAWAVSYTFNFLMSWSSYGTFLIYAINALAI VFI	328	
QY	520	LWVPETKG	527	
Dd	329	AIVPETKG	336	

RESULT 12
 G84864
 probable membrane transporter [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
 C:Accession: G84864
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
 Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: A84420; MUID:20083487; PMID:10617197
 A:Accession: G84864
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-521 <STO>
 A:Cross-references: UNIPROT:O22848; GB:AE002093; NID:g2289003; PIDN:AAB64332.1; GSPDB:GN
 C:Genetics:
 A:Gene: At2g43330
 A:Map position: 2
 C:Superfamily: glucose transport protein

QY	239	FLAYLLGMF---IP--WRLLA	VIGALPCTMLIPGLFFIPESPRWLAKMNLTE	DCETSLQV	293
Db	186	FLSYLVNSAFTQVPGTWRWMLG	VGVPVAVIQFILMLFMPESPRWLFMKNRKAB---	AIQV	242
QY	294	L-RGFE-TDITTEVNDIKRAVASSS--	KRTISFQELNQKKYRTPLLLGILLVLQNL	SG	349
Db	243	LARTYDISRLEDEIDHLSAAEEEEKQ	RKTGVYLDVFRSKELRLAFLAGAGLQAFQ	QFTG	302
QY	350	INGVLFYASSIFKAAGVTNSDLA--	TCSLGAIQVLATGVTTWLLDRAGRILLI	ISTSG	406
Db	303	INTVMYSPPTIQMAGFHSNQLALFL	SLIVAAWNAAGTVVGIYFIDHCGRKKLAL	SSLFG	362
QY	407	MTLCLLAVSVVFFLKDNISQDSNS	YIILTMISLVGIVSFVITFSFCMGAIP	WLMSEILP	466
Db	363	VIISLLILSVSFFKQSETSSDGG	LY---GWLAVLGLALYIVFFAPGMGP	VPTVNSEIYP	419
QY	467	VSIKSLGGSATLANWLTSPAITMT-	TNMLTWSVGGTFLSYMVVSAFTIVFV	VLWVPET	525
Db	420	QQYRGICGGSATVNVISNLIVAQ	FTLTABAAAGTGMTFLILAGIAVLAVI	FVIVFPET	479
QY	526	KG	527		
Db	480	Q	481		

RESULT 13
A31318
glucose transporter-like protein - human
C:Species: Homo sapiens (man)
C:Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 09-Jul-2004
C:Accession: A31318
R:Fukumoto, H.; Seino, S.; Imura, H.; Seino, Y.; Eddy, R.L.; Fukushima, Y.; Byers, M.G.;
Proc. Natl. Acad. Sci. U.S.A. 85, 5434-5438, 1988
A:Title: Sequence, tissue distribution, and chromosomal localization of mRNA encoding a
A:Reference number: A31318; MUID:88289735; PMID:3399500
A:Accession: A31318
A:Molecule type: mRNA
A:Residues: 1-524 <FUK>
A:Cross-references: UNIPROT:P11168; GB:J03810; NID:G187133; PIDN:AAA59514.1; PID:G307125
C:Genetics:
A:Gene: GDB:SLC2A2; GLUT2
A:Cross-references: GDB:119995; OMIM:138160
A:Map position: 3q26.2-3q27
C:Superfamily: glucose transport protein
C:Keywords: transmembrane protein

Query Match	20.7%;	Score 575.5;	DB 2;	Length 524;
Best Local Similarity	33.1%;	Pred. No. 2e-33;		
Matches 136;	Conservative 94;	Mismatches 147;	Indels 34;	Gaps 11;
Qy	132	NLSGSEFSAFGLSNVGGMVGAIASGQMAEYIGRKGLMFAIPNIIGWLAIISPAK----	D	188
Db	98	SLSVSSFA-----VGGMTASFPGGWLGDITLGRKAMLVANILSLVGALLMGFSKLGPS		150
Qy	189	ASFLYMGRLLEGFGVGIISYTPVPVYIAEISPQNMARGALSVNQLSVTFGIF-----	LAY	242
Db	151	HILLIAGRSISGLYCYGLISGLVPMYIGEIAPTALRGALGFPHQLAIVTGILISQIIIGLEF		210
Qy	243	LLGMFIPWRLLAVIGALPCTMLIPGLFFIPESPRL-AKMNLTEDCETSLQVLRGFETDI		301
Db	211	ILGNYDLWHILLGLSGVRALLQSLLLFFCPESPRLYIKLDEEVKAKQSKRLRGYD-DV		269
Qy	302	TTEVNDIK--RAVASSKRTTISFQELNQKKYRTPLLIGICLLVLQNLSGINGVLFYASS		359
Db	270	TKDINEMRKEREASSEQVSI-IQLFTNSSYRQPILVALMLHVAQQFSGINGIFYYSTS		328
Qy	360	IFKAAGVTNSDLATCSLGAIQVLATGVTTWLLDRAGRILLIISTSGMTLCLLAVSVVFF		419
Db	329	IFQTAGISKPPVATIGVGAVMVFTAVSVFLVEKAGRRSLFLIGMSGMFVCAIFMSVGLV		388
Qy	420	LKDNISQDSNSYYILMTISLVGIVSFVITFSFGMGALPWLMMSEILPVSIKSLGGSIA TL		479

Db 389 L-----LNKFSNMSYVSMIAIFLVSFEEIGPGPIPFWMVAEFFSQGPRPAALAAIAAF 441

QY 480 ANWLTSFAITMTTNLMLTWSVGGTFLSYM---VVSFTIVFVVLWVPETKG 527

Db 442 SNWTCNFIVALCFQYIADFC--GPYVFFLFGVLLAFTL-FTFFKVPETKG 489

RESULT 14

T51485

sugar transporter-like protein - Arabidopsis thaliana

N;Alternate names: protein T21H19_70

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 09-Jul-2004

C;Accession: T51485

R;Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.; Mew submitted to the Protein Sequence Database, August 2000

A;Reference number: Z25394

A;Accession: T51485

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-560 <SAT>

A;Cross-references: UNIPROT:Q9LF13; EMBL:AL391148

A;Experimental source: cultivar Columbia; BAC clone T21H19

C;Genetics:

A;Map position: 5

A;Introns: 86/1; 124/2; 161/1; 205/2; 238/3; 276/2; 307/3; 359/1; 408/1; 432/3; 462/2; 5

A;Note: T21H19_70

C;Superfamily: glucose transport protein

Query Match 20.4%; Score 566.5; DB 2; Length 560;

Best Local Similarity 29.4%; Pred. No. 9.4e-33;

Matches 133; Conservative 95; Mismatches 179; Indels 45; Gaps 7;

QY 105 IVALGPIQGTSGFSPTQDAMVRDLNLS-----ISEFSAF 141

Db 111 VACLGAIFGYHLGVNNGALEYLAQDLGIAENTVLOGKYMHIHFFTPVNGMIVSSLLA- 169

QY 142 GSLSNVGMVGAIASGQMAEYIGRKGSMLMIAAIPNIICWLAIISFAKDASFLYMGRLLEGF 201

Db 170 -----GATVGSFTGGALADKFGTRTFQDLDAIPLAIGAFLCATAQSVQTMIVGRLLAGI 223

QY 202 GVGLISYTVPVYIAEISPNMRGALGSVNQLSVTFGIFLAYLLGMFIP-----WRL LAVI 256

Db 224 GIGISSAIVPLYISEISPTIRGALGSVNQLFICIGILAALIAGLPLAANPLWRTWFGV 283

QY 257 GALPCTMLIPGLFFIPESPRWLAKMNLTEDCETSLQVLRGFETDITTEVNDIKRAVASSS 316

Db 284 AVTSPVLLAIGMAFSPESPRWLQVQGVSEAEKAITLYGKER-VVELVRDLS-ASGQGS 341

QY 317 KRTTISFOELNQKRYRTPLLLGIGLLVLQNLSGINGVLFYASSIFKAAGVTNSDLATCSL 376

Db 342 SEPEAGWFDLFSSRYKVVSVGAALFLFQQLAGINAVVYSTSVRSAGIQSDVAASALV 401

QY 377 GAIQVATGVTWLLDRAGRRILLIISTSGMTLCLLAVSVVFFLKDNI SQDSNSYYILTM 436

Db 402 GASNVFGTAVASSLMDKMGKSLLLTSFGGMALSMLLLSFTWKALAAYSGT----- 454

QY 437 ISLVGIYSFVITFSFGMGAIPWLMNMSEILPVSIKSLGGSIATLANWLTSTFAITM-TTNLM 495

Db 455 LAVVGTVLXVLSFSLGAGVPVAPALLPEIFASRIRAKAVALSLGMMHWSNFVIGLYFLSVV 514

QY 496 LTWSVGGTFLSYMVVSFTIVFVVLWVPETKG 527

Db 515 TKFGISSVVLGFAGVCVLAFLYIAGNVVETKG 546

RESULT 15

A41264

glucose transport protein 3 - chicken

C;Species: Gallus gallus (chicken)

C;Date: 28-May-1992 #sequence_revision 28-May-1992 #text_change 04-Sep-1998

C;Accession: A41264

R;White, M.K.; Rall, T.B.; Weber, M.J.

Mol. Cell. Biol. 11, 4448-4454, 1991

A;Title: Differential regulation of glucose transporter isoforms by the src oncogene in A;Reference number: A41264; MUID:91342646; PMID:1875932

A;Accession: A41264

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-496 <WHI>

A;Cross-references: GB:M37785

C;Superfamily: glucose transport protein

C;Keywords: transmembrane protein

Query Match 20.3%; Score 563.5; DB 2; Length 496;

Best Local Similarity 29.6%; Pred. No. 1.3e-32;

Matches 140; Conservative 95; Mismatches 199; Indels 39; Gaps 10;

QY 89 MAVLRSHVSAFLCTLLIVALGPIQGTSGFSPTQDAMVRDLNLSISE-----F 138

Db 1 MADKKITASLIYAVSVAAIGSLQFGYNTGVINAPEKIIQAFYNRTLRSQSGETISPELL 60

QY 139 SAFGSLS---NVGGMVGAIASGQMAEYIGRKGSMLMIAAIPNIIG--WLAIS-FAKDASF 191

Db 61 TSLWSLSVAIFSVMGIGSFSVSLFVNRFGRRNSMLLVNLAFAAGALMALSKIAKAVEM 120

QY 192 LYMRLLLEGFGVGIISYTVPVYIAEISPNMRGALGSVNQLSVTFGIFLAYLLGM----- 246

Db 121 LIIGRFIIGLCGLCTGTFVPMIYSEVSPTSLRGAFGLTNQLGIVVGLVAQIFGLEGIMG 180

QY 247 -FIPWRLAVIGALPCTMLIPGLFFIPESPRWLAKMNLTEDCETSLQVLRGFETDITTE 304

Db 181 TEALWPLLGLFTIVPAVLQCVALLFCPESPRFLINKMEEHKAQTVLQKLRGTQ-DVSQD 239

QY 305 VNDIKRAVASSSKRTTISFOEL-NQKKYRTPLLLGIGLLVLQNLSGINGVLFYASSIFKA 363

Db 240 ISEMKEESAKMSQEKATVLELERSPNYRQPIIISITLQSLSGINAVFYYSTGIFER 299

QY 364 AGVTNSDLATCSLGAIOVLATGVTWLLDRAGRRILLIISTSGMTLCLLAVSVVFFLXDN 423

Db 300 AGITQPVYATIGAGVNTVFTVVSLELVERAGRRTLHLVGLGMVCAAVMTIALALKEK 359

QY 424 ISQDSNSYYILTMISLVGIVSFVITFSFGMGAIPWLMNMSEILPVSIKSLGGSIATLANWL 483

Db 360 -----WIRYISIVATFGFVALFEIGPGPIPFWIFVAELFSQGPRAAMAVAGCSNWT 410

QY 484 TSFAITMTTNLMLTWSVGGTFLSYMVVSFTIVFVVLWVPETKGXN----SRG 532

Db 411 SNFLVGMFLPPYAEKLCGPYVFLIFLVFLLIFFIFTYFKVPETKGRTFEDISRG 463

Search completed: October 13, 2004, 11:39:11

Job time : 16.714 secs

is Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 13, 2004, 11:40:19 ; Search time 4488.99 Seconds
(without alignments)
4489.025 Million cell updates/sec

Title: US-10-051-909-36
Perfect score: 2779
Sequence: 1 PSSSSSFRPAGKKKKKKNQG.....TIFVSLSIQRQLQWLPECLS 553

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 32822875 seqs, 18219865908 residues
Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2 1/USPTO.spool.h/US10051909/runat 13102004 123336 19917/app_query.fasta_1.1678
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosu62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10051909 @CGN 1 1 6263 @runat 13102004 123336 19917 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1733.5	62.4	1837	3	CNS0ACUC
2	1219	43.9	854	7	CN125032
3	1102	39.7	868	6	CB675064
4	1065	38.3	682	6	CA191028
5	1021	36.7	635	6	CA143053
6	1004.5	36.1	764	6	CB648359
7	984.5	35.4	1155	7	CK211005
8	984	35.4	674	1	AV939950
9	976	35.1	714	6	CA262285

10	975	35.1	681	1	AV913420
11	962.5	34.6	718	7	CR291514
12	962.5	34.6	762	1	AJ794429
13	932	33.5	701	6	CD871121
14	930.5	33.5	1675	3	CNS0A8HU
15	925	33.3	769	4	BI933646
16	919.5	33.1	1737	3	CNS09ZVC
17	914	32.9	772	6	CB683315
18	907.5	32.7	1666	3	CNS0A8HV
19	907	32.6	666	5	BQ766951
20	904.5	32.5	627	6	CA180355
21	904	32.5	652	6	CA201877
22	902	32.5	716	7	CF451024
23	901.5	32.4	748	6	CB347677
24	896.5	32.3	1601	3	CNS0A620
25	896	32.2	685	5	BQ862009
26	893	32.1	566	4	BM325827
27	890	32.0	791	6	CB892803
28	887.5	31.9	746	6	CD483180
29	882	31.7	594	6	CD203945
30	882	31.7	772	7	CF835377
31	870.5	31.3	744	1	AJ796408
32	869	31.3	641	6	CA246696
33	864.5	31.1	927	7	CK153511
34	859	30.9	548	6	CA143695
35	850	30.6	627	4	BJ249193
36	847	30.5	663	7	CN906216
37	845	30.4	670	5	BQ855360
38	838	30.2	595	2	BE599181
39	830.5	29.9	651	4	BG522368
40	828	29.8	667	6	CD893209
41	827	29.8	626	7	CN904481
42	825	29.7	656	5	BQ913239
43	822	29.6	1054	3	AY111571
44	821	29.5	686	7	COL16723
45	820.5	29.5	714	6	CA262252

ALIGNMENTS

RESULT 1
CNS0ACUC 1837 bp mRNA linear HTC 06-FEB-2004
LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone
DEFINITION GSLTSL242B10 of Adult vegetative tissue of strain col-0 of
Arabidopsis thaliana (thale cress).
ACCESSION BX815012
VERSION BX815012.1 GI:42472886
KEYWORDS HTC; GSLT cDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE 1 (bases 1 to 1837)
AUTHORS Castelli,V., Aury,J.M., Jallion,O., Wincker,P., Clepet,C.,
Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,
Temple,G., Caboche,M., Weissbach,J. and Salanoubat,M.
TITLE Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1837)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C.,

Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
length
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.

FEATURES
source

Location/Qualifiers
1..1837
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Col-0"
/db_xref="taxon:3702"
/clone="GSLTSL242B10"
/tissue_type="Adult vegetative tissue"
/plasmid="pCMVSPORT_6"
complement(1..1837)
/gene="At1g19450"

gene

ORIGIN

Alignment Scores:

Pred. No.: 1.6e-168 Length: 1837
Score: 1733.50 Matches: 338
Percent Similarity: 83.23% Conservative: 69
Best Local Similarity: 69.12% Mismatches: 74
Query Match: 62.38% Indels: 9
DB: 3 Gaps: 2

US-10-051-909-36 (1-553) x CNSOACUK (1-1837)

QY	52	GluGluSerGlySerAspHisAspGlyValLeuArgArgProLeuLeuAsnThrGlySer	71
DB	213	GAAGAGGGAAGAAATGAT-----CTCGCGCACCCTTCTTACACACAGGAAGT	260
QY	72	TrpTyrArgMetSerSerArgGlnSerSerPheAlaProGlyThrSerSerMetAlaVal	91
DB	261	TGGTATCGGATGGGTCGAGACAATCTAGTATG-----TTGGAATCGTCTCAAGTT	311
QY	92	LeuArgGluSerHisValSerAlaPheLeuCysThrLeuIleValAlaLeuGlyProIle	111
DB	312	ATTCGAGACAGCTCAATTTCTGTCTTAGCTTGTTGATTTGTTGCTTGTGCTCTATT	371
QY	112	GlnPheGlyPheThrSerGlyPheSerSerProThrGlnAspAlaMetValArgAspLeu	131
DB	372	CAATTCGGATTTCACCTTGTTGTTATTCCTCACCACACTCAAGTCGCCATCTACTAAGGATCTT	431
QY	132	AsnLeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyGlyMetVal	151
DB	432	GTTTAACTGTATCAGAGTACTCTGTGTTGGGCTCTCTATCCAAATGGGTGCTATGGTT	491
QY	152	GlyAlaIleAlaSerGlyGlnMetAlaGluTyrIleGlyArgLysGlySerLeuMetIle	171
DB	492	GGAGCAATTGCGAGTGGTCAGATTGCTGAATACCTTGGACGGAAAGGGTCTCTGATGATT	551
QY	172	AlaAlaIleProAsnIleIleGlyTrpLeuAlaIleSerPheAlaLysAspAlaSerPhe	191
DB	552	GCTGCAATTCCCAATATATTGGATGGCTTTTCGATATCATTTTCCGAAAGATACTTCTTTT	611
QY	192	LeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTyrThrValPro	211
DB	612	CTTTACATGGGAAGATTGTTAGAAGGCTTCGGAGTTGGGATTATCTCTATACGGTTCCC	671
QY	212	ValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGln	231
DB	672	GTATATATAGCTGAGATCGTCCACAGACCACATGAGAGGAGCCTTAGGTTCAAGTTAACCCAG	731
QY	232	LeuSerValThrPheGlyIlePheLeuAlaTyrLeuLeuGlyMetPheIleProTyrArg	251
DB	732	CTTTCTGTAAACAATTGGGATAATCTGGCGTATTTACTCGGTCTCTTTGTTCCATGGAGA	791
QY	252	LeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPheIle	271

DB	792	ATTCTTGCAGTTTTGGAGTATTGCCAATGTACATTGTTGATACCGGGTCTTTTTCATT	851
QY	272	ProGluSerProArgTrpLeuAlaLysMetAsnLeuThrGluAspCysGluThrSerLeu	291
DB	852	CCTGAATCTCCTCGGTGGTGGCAAGATGGGTTTGACAGATGATTCGAAACTTCATTG	911
QY	292	GlnValLeuArgGlyPheGluThrAspIleThrGluValAsnAsp-IleLysArgAl	311
DB	912	CAAGTTCTTCGTGGATTGAGACTGATATACCGTTGAGGTTAATGAAATCAAGAGATC	971
QY	311	aValAlaSerSerSerLysArgThrThrIleSerPheGlnGluLeuAsnGlnLysLysTy	331
DB	972	TGTGGCATCATCTAGCAAAACGTTCTGCAGTTCCGTTTGTAGACCTCAAGCGCAGGATA	1031
QY	331	rArgThrProLeuLeuGlyIleGlyLeuLeuValLeuGlnAsnLeuSerGlyIleAs	351
DB	1032	CTATTCCCACATGATGGTGGTATAGGGCTGCTGCACCTTCAACAACCTCGAGGAATCAA	1091
QY	351	nGlyValLeuPheTyrAlaSerSerIlePheLysAlaAlaGlyValThrAsnSerAspLe	371
DB	1092	TGGTGTCTTGTCTATTTCGAGTACAATTTTGAATCTGCAGGGGTGACATCAAGTAATGT	1151
QY	371	uAlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThrGlyValThrTrpLeuLe	391
DB	1152	GGCGACATTTGGAGTTGGCGTTGTTTCAGGTAGTGGGACGTGGATAGCGACATGGTTGGT	1211
QY	391	uAspArgAlaGlyArgIleLeuLeuIleIleSerThrSerGlyMetThrLeuCysLe	411
DB	1212	GGATAAGCAGGTCGTCGGCTTCTACTCATGATCTCTCTATCGGAATGACGATAAGCCT	1271
QY	411	uLeuAlaValSerValPhePheLeuLysAspAsnIleSerGlnAspSerAsnSerTy	431
DB	1272	CGTGATTGTAGCAGTTGCATTTTACCTTAAGGAATTTGTATCACCTGATTCACACATGTA	1331
QY	431	rTyrIleLeuThrMetIleSerLeuValGlyIleValSerPheValIleThrPheSerPh	451
DB	1332	CAACATTTCTAAGCATGGTTTCCGTAGTTGGAGTTGTGGCTATGGTTATCTCTGCTCTCT	1391
QY	451	eGlyMetGlyAlaIleProTrpLeuMetMetSerGluIleLeuProValSerIleLysSe	471
DB	1392	AGGAATGGGACCAATTCATGGCTGATTATGTCTGAGATTCGCCAGTGAATATAAGGG	1451
QY	471	rLeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeuThrSerPheAlaIleThrMetTh	491
DB	1452	TTTAGCCGGAAGTATAGCAACTTTGTTAACTGGTTCGTGTCTAGTGTAGTACATGAC	1511
QY	491	rThrAsnLeuMetLeuThrTrpSerValGlyGlyThrPheLeuSerTyrMetValValse	511
DB	1512	TGCAAAATATGCTGTAGCATGGAGCAGTGGGGGAACTTTCACTCTCTACGCTTGGTTTG	1571
QY	511	rAlaPheThrIleValPheValValLeuTrpValProGluThrLysGly***AsnSerAr	531
DB	1572	TGGATTCACCTGTGGTTTGTGAGTCTTTGGGTTCTCTGAGACTAAAGGAAA-AACGCTTG	1630
QY	531	gGlyAspThrIlePheValSerLeu	539
DB	1631	AAGAGATCCAAGCTTTGTTTCAGATG	1655

RESULT 2

CN125032- LOCUS	854 bp	mRNA	linear	EST 01-APR-2004
DEFINITION	RHOH1_8_A07.g1_A002 Acid- and alkaline-treated roots Sorghum			
ACCESSION	bicolor cDNA clone RHOH1_8_A07_A002 5', mRNA sequence.			
VERSION	CN125032.1			
KEYWORDS	EST.			
SOURCE	Sorghum bicolor (sorghum)			
ORGANISM	Sorghum bicolor			
REFERENCE	1 (bases 1 to 854)			
AUTHORS	Cordonnier-Pratt,M.-M., Suzuki,Y., Sugano,S., Klein,R.R., Liang,C.,			


```
/tissue_type="Leaf"  
/dev_stage="3 week"  
/lab_host="DH10B"  
/clone_lib="OSJNEE"  
/note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:  
XhoI; 24 hrs after inoculation with Rice Blast (70-15)"
```

ORIGIN

Alignment Scores:
Pred. No.: 1.93e-103 Length: 868
Score: 1102.00 Matches: 208
Percent Similarity: 87.89% Conservative: 46
Best Local Similarity: 71.97% Mismatches: 35
Query Match: 39.65% Indels: 0
DB: 6 Gaps: 0

US-10-051-909-36 (1-553) x CB675064 (1-868)

QY 206 IleSerThrValProValThrIleAlaGluIleSerProGlnAsnMetArgGlyAla 225
Db 1 ATATCCTATGTGGTCCAGTTTATATAGCAAAATGCTCCACAAACAATGAGAGGAGCT 60
QY 226 LeuGlySerValAsnGlnLeuSerValThrPheGlyIlePheLeuAlaTyrLeuLeuGly 245
Db 61 CTGGTTTCAGTCAATCAGCTTTCTGTTACTATTGGCATATTGCTTGCCTACTTGTAGGC 120
QY 246 MetPheIleProTrpArgLeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIle 265
Db 121 ATGTTTGTCCCTGGAGAAATCTCTCCGTTCTCGGTATTTTACCATTGTTCAATCTGATA 180
QY 266 ProGlyLeuPhePheIleProGluSerProArgTrpLeuAlaLysMetAsnLeuThrGlu 285
Db 181 CCTGGATTGTTCTTTATCCCTGAATCACCAGGTGGCTGGCAAAATGGGAAAGATGAAG 240
QY 286 AspCysGluThrSerLeuGlnValLeuArgGlyPheGluThrAspIleThrThrGluVal 305
Db 241 GATTTTGAATCCTCGCTGCAAGTATTGCGAGGATTTGAAACAGACATCGCTGTAGAAGTA 300
QY 306 AsnAspIleLysArgAlaValAlaSerSerLysArgThrThrIleSerPheGlnGlu 325
Db 301 AATGAATAAAGAGACAGTTCAATCATCAGGAGGAGGACGACATACGATTTGCAGAT 360
QY 326 LeuAsnGlnLysLysTyrArgThrProLeuLeuLeuGlyIleGlyLeuValLeuGln 345
Db 361 ATCAAGCAGAGAGATAGTAGTACCTCTAATGATAGGAATTGGTCTCTTGTACTGCAG 420
QY 346 AsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSerIlePheLysAlaAlaGly 365
Db 421 CAGTTAAGTGGTGTGAATGCAATCTATTTATGCTGTAGTATCTTCAAAGCCCGCGGT 480
QY 366 ValThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThrGly 385
Db 481 CTACGAATAGTAATCTAGCAACATTTGGTCTGGGGTTGTTTCAGGTGGTGTCTACTGGA 540
QY 386 ValThrThrTrpLeuLeuAspArgAlaGlyArgArgIleLeuLeuIleSerThrSer 405
Db 541 GTGACAACTTGGTTGACTGACAAAGCTGGTCGACGACTCCTCCTCATATCTCTACTACA 600
QY 406 GlyMetThrLeuCysLeuLeuAlaValSerValValPhePheLeuLysAspAsnIleSer 425
Db 601 GGGATGACCATTAATCTGTTGTTGTTCTGTCATTTTGTGAAGCACAACATAACT 660
QY 426 GlnAspSerAsnSerTyrIleLeuThrMetIleSerLeuValGlyIleValSerPhe 445
Db 661 AATGGTTCTCATTTATCTGTAATGATGATGCTTTGCTGGTGGGCTTGTGGCATTT 720
QY 446 ValIleThrPheSerPheGlyMetGlyAlaIleProTrpLeuMetMetSerGluIleLeu 465
Db 721 GTGATTTTCATTTCTCTGTTGGGAGCCCATACCATGGATCATATGTTCTGAGATTCTT 780
QY 466 ProValSerIleLysSerLeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeuThrSer 485
Db 781 CCTGTAAATATCAAGAGCCTTGTCTGGAGCGTGTGCAACCCCTTGCAAACTGGTTGACGGCT 840

QY 486 PheAlaIleThrMetThrThrAsnLeu 494
Db 841 TGGCTCATTACGATGACAGCAAGCTTG 867

RESULT 4

CA191028

LOCUS

DEFINITION

5', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Saccharum officinarum

Saccharum officinarum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum

complex.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 682)

Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.

The libraries that made SUCEST

Genet. Mol. Biol. 24 (1-4), 1-7 (2001)

Contact: Arruda P

Centro de Biologia Molecular e Engenharia Genetica

Universidade Estadual de Campinas

Caixa Postal 6010, 13083-970, Campinas SP, Brazil

Tel: 55 19 3788 1137

Fax: 55 19 3788 1089

Email: parruda@unicamp.br

Clone distribution: clone distribution information can be found

through the Brazilian Clone Collection Center (BCCC) at

http://www.bcccenter.fcav.unesp.br

Plate: C03 row: A column: 08

Seq primer: T7 Promoter Primer.

Location/Qualifiers

1..682

/organism="Saccharum officinarum"

/mol_type="mRNA"

/db_xref="taxon:4547"

/clone="SCCCT2C03A08"

/lab_host="DH10B"

/clone_lib="RT2"

/note="Organ: Root tips(0.3cm-long) from adult plants;

Vector: pSport1; Site_1: SalI; Site_2: NotI; An

unidirectional cDNA library generated from [Root

tips(0.3cm-long) from adult plants]. cDNA was prepared

from polyA+ mRNA using Superscript Plasmid System Kit

(Invitrogen). The double-strand cDNAs were fractionated

in a sepharose CL-2B 40cm-columns and fragments sizing

between 0.8 and 1.5 Kb were directionally cloned into the

vector. Details of each source of RNA and library

construction can be obtained at

http://sucest.lad.ic.unicamp.br/public"

ORIGIN

Alignment Scores:

Pred. No.: 9.18e-100 Length: 682

Score: 1065.00 Matches: 218

Percent Similarity: 97.36% Conservative: 3

Best Local Similarity: 96.04% Mismatches: 6

Query Match: 38.32% Indels: 1

DB: 6 Gaps: 0

US-10-051-909-36 (1-553) x CA191028 (1-682)

QY 256 IleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPhePheIleProGluSerPro 275

Db 2 ATCGGAGCCTTGCCCTGCACAGTGTGTGATTCTCTGGACTATTCTTCAATCCAGAAATCTCCC 61

QY 276 ArgTrpLeuAlaLysMetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuArg 295

Db 62 ACATGGCTGGCAAGATGAATTTGATGGAAGATTGCGAGACGTCCTTCAAGTCTGAGG 121


```
QY 296 GlyPheGluThrAspIleThrThrGluValAsnAspIleLysArgAlaValAlaSerSer 315
    |||||
Db 122 GGGTTTGAGACTGC-ATCACGACAGAAAGTGAATGATATAAAGAGGGCAGTGACATCATCA 180
    |||||

QY 316 SerLysArgThrThrIleSerPheGlnGlnGluLeuAsnGlnLysLysTyrArgThrProLeu 335
    |||||
Db 181 AGTAAGAGGACTACAAATCAGTTTCAAGAAATTAACCAAAAGAAATACCGCAGCGGCTA 240
    |||||

QY 336 LeuLeuGlyIleGlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPhe 355
    |||||
Db 241 CTTCTAGGAATTGGCCTACTTGCTACTGCAAAATCTAAGTGAATCAACGGTGCTGTTT 300
    |||||

QY 356 TyrAlaSerSerIlePheLysAlaAlaGlyValThrAsnSerAspLeuAlaThrCysSer 375
    |||||
Db 301 TATGCAAGTAGCATCTTCAAAGCTGCAGGTGTTACAAACAGCGACTTGGCCACCTGTTCA 360
    |||||

QY 376 LeuGlyAlaIleGlnValLeuAlaThrGlyValThrThrTrpLeuLeuAspArgAlaGly 395
    |||||
Db 361 CTTGGAGCTATCCAGGTCCTTGCTACTGGAGTTACGACATGGTTGTTAGACAGAGCTGA 420
    |||||

QY 396 ArgArgIleLeuLeuIleIleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSer 415
    |||||
Db 421 CGACGCATGCTTCTCATATTATTTCTACCTCTGGCATGACTCTATGCCTTCTTCGCGTTTCT 480
    |||||

QY 416 ValValPhePheLeuLysAspAsnIleSerGlnAspSerAsnSerTyrTyrIleLeuThr 435
    |||||
Db 481 GTTGATATTTTCTCAAGGATAAGATTTCACAGGAATCTAACTCGTACTACATCTTAAC 540
    |||||

QY 436 MetIleSerLeuValGlyIleValSerPheValIleThrPheSerPheGlyMetGlyAla 455
    |||||
Db 541 ATGATCTCCTTGGTTGCTATCGTGGCTTTTGTCAATACCTTCTCGTTTGGTATGGGTGCC 600
    |||||

QY 456 IleProTrpLeuMetMetSerGluIleLeuProValSerIleLysSerLeuGlyGlySer 475
    |||||
Db 601 ATTCCATGGCTCATGATGCTCAGATCCTTCCAGTTAGCATCAAGAGTCTCGGCGGAAGC 660
    |||||

QY 476 IleAlaThrLeuAlaAsnTrp 482
    |||||
Db 661 ATCGGACACTGGCCAAATGG 681
    |||||

RESULT 5
CA143053
LOCUS CA143053 635 bp mRNA linear EST 24-SEP-2003
DEFINITION SCQSR2032A08.g RT2 Saccharum officinarum cDNA clone SCQSR2032A08
5', mRNA sequence.
ACCESSION CA143053
VERSION CA143053.1 GI:35038703
KEYWORDS EST.
SOURCE Saccharum officinarum
ORGANISM Saccharum officinarum
REFERENCE 1 (bases 1 to 635)
AUTHORS Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
TITLE The libraries that made SUCEST
JOURNAL Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
COMMENT Contact: Arruda P
Centro de Biologia Molecular e Engenhariaia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bcccenter.fcav.unesp.br
Plate: 032 row: A column: 08
Seq primer: T7 Promoter Primer.
Location/Qualifiers
source 1..635
```

```
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCQSR2032A08"
/lab_host="DH10B"
/clone_lib="RT2"
/note="Organ: Root tips(0.3cm-long) from adult plants;
Vector: pSport1; Site 1: SalI; Site 2: NotI; An
unidirectional cDNA library generated from [Root
tips(0.3cm-long) from adult plants]. cDNA was prepared
from polyA+ mRNA using SuperScript Plasmid System Kit
(Invitrogen). The double-strand cDNAs were fractionated
in a sepharose CL-2B 40cm-columns and fragments sizing
between 0.8 and 1.5 Kb were directionally cloned into the
vector. Details of each source of RNA and library
construction can be obtained at
http://sucest.lad.ic.unicamp.br/public"
```

ORIGIN

```
Alignment Scores: 3.05e-95 Length: 635
Pred. No.: 1021.00 Matches: 204
Score: 1021.00
Percent Similarity: 98.10% Conservative: 3
Best Local Similarity: 96.68% Mismatches: 4
Query Match: 36.74% Indels: 0
DB: 6 Gaps: 0

US-10-051-909-36 (1-553) x CA143053 (1-635)

QY 256 IleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPhePheIleProGluSerPro 275
    |||||
Db 2 ATCGGAGCCTTGGCCTGCACAGTGTGATTCCTGGACTATTCTTCATTCAGAAATCTCCC 61
    |||||

QY 276 ArgTrpLeuAlaLysMetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuArg 295
    |||||
Db 62 AGATGGCTGGCAAAGATGAATTTGATGGAAGATTGGAGACGTCCCTACAAGTGTGAGG 121
    |||||

QY 296 GlyPheGluThrAspIleThrThrGluValAsnAspIleLysArgAlaValAlaSerSer 315
    |||||
Db 122 GGGTTTGAGACTGACATCACGACAGAAAGTGAATGATATAAAGAGGGCAGTGACATCATCA 181
    |||||

QY 316 SerLysArgThrThrIleSerPheGlnGlnGluLeuAsnGlnLysLysTyrArgThrProLeu 335
    |||||
Db 182 AGTAAGAGGACTACAATCAGTTTTCAAGAAATTAACCAAAAGAAATACCGCAGCGCTA 241
    |||||

QY 336 LeuLeuGlyIleGlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPhe 355
    |||||
Db 242 CTTCTAGGAATTGGCCTACTTGCTACTGCAAAATCTAAGTGAATCAACGGTGCTGTTT 301
    |||||

QY 356 TyrAlaSerSerIlePheLysAlaAlaGlyValThrAsnSerAspLeuAlaThrCysSer 375
    |||||
Db 302 TATGCAAGTAGCATCTTCAAAGCTGCAGGTGTTACAAACAGCGACTTGGCCACCTGTTCA 361
    |||||

QY 376 LeuGlyAlaIleGlnValLeuAlaThrGlyValThrThrTrpLeuLeuAspArgAlaGly 395
    |||||
Db 362 CTTGGAGCTATCCAGGTCCTTGCTACTGGAGTTACGACATGGTTGTTAGACAGAGCTGA 421
    |||||

QY 396 ArgArgIleLeuLeuIleIleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSer 415
    |||||
Db 422 CGACGCATGCTTCTCATATTATTTCTACCTCTGGCATGACTCTATGCCTTCTTGGCGTTTCT 481
    |||||

QY 416 ValValPhePheLeuLysAspAsnIleSerGlnAspSerAsnSerTyrTyrIleLeuThr 435
    |||||
Db 482 GTTGTATTTTCTCAAGGATAAGATTTCACAGGATTCTAAGTCTGCTACTACATCTTAACT 541
    |||||

QY 436 MetIleSerLeuValGlyIleValSerPheValIleThrPheSerPheGlyMetGlyAla 455
    |||||
Db 542 ATGATCTCCTTGGTTGCTATCGTGGCTTTTGTTCATTACCTTCTCGTTTGGTATGGGTGCC 601
    |||||

QY 456 IleProTrpLeuMetMetSerGluIleLeuPro 466
    |||||
Db 602 ATTCCATGGCTCATGATGCTGAGATCCTTCCA 634
    |||||
```

```
RESULT 6
CB648359
LOCUS
DEFINITION
OSJNEB11L10.f OSJNEB Oryza sativa (japonica cultivar-group) cDNA
clone OSJNEB11L10 5', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE
1 (bases 1 to 764)
Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea
Unpublished (2003)
JOURNAL
COMMENT
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
FORWARD: gta aaa cga cgg cca gtcg
BACKWARD: gga aac agc tat gac cat g
Plate: 11 row: L column: 10
Seq primer: gta aaa cga cgg cca gtcg.
Location/Qualifiers
1..764
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="OSJNEB11L10"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSJNEB"
/note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
XhoI; 24 hrs after inoculation with Rice Blast (Che
86061)"
ORIGIN
Alignment Scores:
Pred. No.: 2.12e-93 Length: 764
Score: 1004.50 Matches: 202
Percent Similarity: 87.40% Conservative: 13
Best Local Similarity: 82.11% Mismatches: 24
Query Match: 36.15% Indels: 7
DB: 6 Gaps: 2
US-10-051-909-36 (1-553) x CB648359 (1-764)
QY 28 GlyArgProAlaSerGluLeuArgThrArgValMetGlyGlySerAsnArgGlyGly 47
Db 35 GGAGCCGAGCGGAGCGAGCGGCGGCGGATGAACGGCGCGGAGCAGCGGGAGGT 94
QY 48 AlaGlyAlaGlyGluGluSerGlySerAspHisAsp-----GlyValLeuArg 63
Db 95 GCGTCGTGCGCGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 154
QY 64 ArgProLeuLeuAsnThrGlySerTrpTyArgMetSerSerArgGlnSerSerPheAla 83
Db 155 AAGCCGCTGCTGAACACGGGGAGCTGGTACAGGATGGGATCGCGGTCCAGCCTCGCCGCC 214
QY 84 ProGlyThrSerSerMetAlaValLeuArgGluSerHisValSerAlaPheLeuCysThr 103
Db 215 -----TCTTCATGGCCGCCCATCCGGGAGTCCCGCTCTCCGCTTCTCTGCACG 265
```

```
QY 104 LeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGlyPheSerSerProThr 123
Db 266 CTCATCGTCGCGCTCGGCCCAATCCAAATTCGGATTCAACAGCGGCTTCTCTCACCCACC 325
QY 124 GlnAspAlaMetValArgAspLeuAsnLeuSerIleSerGluPheSerAlaPheGlySer 143
Db 326 CAGGACGCCATCATCCGCGACCTCAAGCTCTCCATCTCCGAGTTCTCGGTTTCG 385
QY 144 LeuSerAsnValGlyGlyMetValGlyAlaIleAlaSerGlyGlnMetAlaGluTyIle 163
Db 386 CTGTCCAACGTCGCGGCCCATGGTCGGAGCGATGCCAGTGGGCGAGATGGCGGAGTACATT 445
QY 164 GlyArgLysGlySerLeuMetIleAlaAlaIleProAsnIleIleGlyTrpLeuAlaIle 183
Db 446 GGCCGGAAGGTCGTGATAATTGCGCGGTTCTTAACATCATTTGGTTGGCTTGCATC 505
QY 184 SerPheAlaLysAspAlaSerPheLeuTyMetGlyArgLeuLeuGluGlyPheGlyVal 203
Db 506 TCCTTTGCAAAAGACGCGTCATTTTATATACATGGGACGCTTGCTGAAGGTTTGGTGT 565
QY 204 GlyIleIleSerTyThrValProValTyIleAlaGluIleSerProGlnAsnMetArg 223
Db 566 GGTGTATATCATATACGGTGCCAGTATACATAGCAGATATCTCATCAGAACACAAGA 625
QY 224 GlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIlePheLeuAlaTyIle 243
Db 626 GGAGCACTTGGCTCGTGAAACAGTTGTCGGTTACCATTTGGTATCTTGTGGCCTATTG 685
QY 244 LeuGlyMetPheIleProTrpArgLeuLeuAlaValIleGlyAlaLeuProCysThrMet 263
Db 686 CTAGGCATGTTTGTCTTGGAGGCTGCTTGCAGTGATAGGAAGCATCCCATGTACATTG 745
QY 264 LeuIleProGlyLeuPhe 269
Db 746 TTAATACCTGGTCTATTTC 763
RESULT 7
CK211005 1155 bp mRNA linear EST 08-DEC-2003
FGAS022833 Triticum aestivum FGAS: Library 5 GATE 7 Triticum
aestivum cDNA, mRNA sequence.
CK211005
CK211005.1 GI:39573395
EST.
Triticum aestivum (bread wheat)
ORGANISM
Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.
REFERENCE
1 (bases 1 to 1155)
AUTHORS
Allard,F., Crosby,W.L., Danyluk,J., Eudes,F., Frick,M., Gaudet,D.,
Genswein,B., Graf,R., Gulick,P., Hrycan,L.D., Laroche,A.,
Links,M.G., McCarthy,E.L., Monroy,A., Muzak,I., Nilson,D.,
Penniket,C., Roach,J.L. and Sarhan,F.
Functional Genomics of Abiotic Stress In Wheat and Canola Crops
Unpublished (2003)
JOURNAL
COMMENT
Contact: Wm L Crosby
Bioinformatics
University of Saskatchewan, Department of Computer Science
1C101 Engineering Building, 57 Campus Drive, Saskatoon,
Saskatchewan, S7N 5A9, Canada
Tel: 306 966 1769
Fax: 306 966 2033
Email: fgas_est@cs.usask.ca
This sequence is the direct result of the Base calling software
Phred (default parameters). It is the raw base calls. To aid in the
identification of the high quality insert the software Lucy
(default parameters) has been run on this sequence. Lucy identified
the region [44,808].
Plate: L5B024 row: L column: 10.
Location/Qualifiers
1..1155
```

FEATURES source

/organism="Triticum aestivum"
/mol_type="mRNA"
/db_xref="taxon:4565"
/clone_lib="Triticum aestivum FGAS: Library 5 GATE 7"
/note="Vector: PCMV.SPORT6; Crown and developmental stages of spike formation in wheat cultivar Norstar. 4 mRNA populations were combined before constructing the library. The first mRNA population is from 1cm crown sections after 30 days of cold acclimation. The second is from 1cm crown sections after 11 days of deacclimation (before deacclimation plants were fully vernalized for 49 days). The third is from different developmental stages of spike formation (5 to 50mm) that still have not emerged from the leaf (dissection required). The last is from different developmental stages of spike and seed formation after having emerged from the leaf (visible). First strand synthesis in this library was done in the presence of methylated dCTP thereby protecting from internal cleavage with NotI."

ORIGIN

Alignment Scores:
Pred. No.: 4.88e-91 Length: 1155
Score: 984.50 Matches: 195
Percent Similarity: 82.83% Conservative: 51
Best Local Similarity: 65.66% Mismatches: 51
Query Match: 35.43% Indels: 3
DB: 7 Gaps: 0

US-10-051-909-36 (1-553) x CK211005 (1-1155)

QY 244 LeuGlyMetPheIleProTrpArgLeuLeuAlaValIleGlyAlaLeuProCysThrMet 263
Db 54 GTAGGCATGTTTGTTCCTCGGAGATCTCTCAGTCCTGGCATTTTACCTGCTCAATC 113
QY 264 LeuIleProGlyLeuPheIleProGluSerProArgTrpLeuAlaLysMetAsnLeu 283
Db 114 CTGATACCTGGTGTTCCTTCATCCCCGAATCACCAGGTGGCTGGCAAAATGGGAAG 173
QY 284 ThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPheGluThrAspIleThrThr 303
Db 174 ATGAGGATTTTGAATCTTCCTGTCAGGTTCTCGCGGATTTGAACTGATATCTCAGCA 233
QY 304 GluValAsnAspIleLysArgAlaValAlaSerSerSerLysArgThrThrIleSerPhe 323
Db 234 GAAGTAAATGAATAAGAGATCAGTTGCCTCATCTAGGAGGAGCAACCATACGATTT 293
QY 324 GlnGluLeuAsnGlnLysLysTyrArgThrProLeuLeuGlyIleGlyLeuLeuVal 343
Db 294 GCAGAGATCAAAACACAGAGATATAGTGTTCCTCTTATGATAGGAATTGGTCTCTTATA 353
QY 344 LeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSerIlePheLysAla 363
Db 354 CTTTCAGCAACTAAGTGGTGTCAATGGCATCTTCTTCTATGCTGCAAGTATCTTCAAAGCT 413
QY 364 AlaGlyValThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIleGlnValLeuAla 383
Db 414 GCTGCTCTTAAATAATAGTAACTAGCAACATGTGTTTGGGGCTGTTTCAGGTGGTTGCT 473
QY 384 ThrGlyValThrThrTrpLeuLeuAspArgAlaGlyArgIleLeuLeuIleIleSer 403
Db 474 ACTGGAATCACAACTGGTTGACTGACAAAGCTGGTCGACGCTACTTCTCATTATCTCT 533
QY 404 ThrSerGlyMetThrLeuCysLeuLeuAlaValSerValPhePheLeuLysAspAsn 423
Db 534 GCTGTAGGATGACAGTCAGTCTTCTCGTTGTTTCTGTGTCATTTGTTGAAGGAAAC 593
QY 424 IleSerGlnAspSerAsnSerTyrTyrIleLeuThrMetIleSerLeuValGlyIleVal 443
Db 594 ATAGAGGAAGCTTCTCATTTACACTCTGTGATGATGATGCTTTCACCTGGCTGGACITGTG 653
QY 444 SerPheValIleThrPheSerPheGlyMetGlyAlaIleProTrpLeuMetMetSerGlu 463

Db 654 GCATTGTGCTTGCATTTTCCCTTGGCATGGGAGCCATCCCATGATCATAATGCTGTGAG 713
QY 464 IleLeuProValSerIleLysSerLeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeu 483
Db 714 ATCTTCTCTGTTAATCATCAAGAGTCTCCCGGAAGCACCCCTCGCAAACTGGATG 773
QY 484 ThrSerPheAlaIleThrMetThrThrAsnLeuMetLeuThrTrpSerValGlyGlyThr 503
Db 774 ACGTCTGGCTCATCACGATACCGGGAGCTTGATGTAACTGGAGCAACGGAGGAACC 833
QY 504 PheLeuSerTyrMetValValSerAlaPheThrIleValPheValLeuTrpValPro 523
Db 834 TTGCTATA-TTCGCGCGGGTGTCCATGGGCACCCCTCTCTGCTGTGCTGTGCGTGC 892
QY 524 GluThrLysGly***AsnSerArgGlyAspThrIlePheValSerLeuSer 540
Db 893 GAGACCAAGGAAG-AACATTGAGGAAA-TCGGTTTTCGTTCCGCTGAAC 941

RESULT 8

AV939950 674 bp mRNA linear EST 18-JAN-2002
LOCUS AV939950 K. Sato unpublished cDNA library, strain H602 adult,
DEFINITION heading stage top three leaves Hordeum vulgare subsp. spontaneum
cDNA clone bah24m12 5', mRNA sequence.

ACCESSION AV939950
VERSION AV939950.1 GI:18235747
KEYWORDS EST.
SOURCE Hordeum vulgare subsp. spontaneum
ORGANISM Hordeum vulgare subsp. spontaneum

REFERENCE 1 (bases 1 to 674)
AUTHORS Sato,K., Saisho,D. and Takeda,K.
TITLE Barley EST sequencing project in NIG and Okayama Univ
JOURNAL Unpublished (2002)
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

FEATURES
source
1..674
/organism="Hordeum vulgare subsp. spontaneum"
/mol_type="mRNA"
/strain="H602"
/sub_species="spontaneum"
/db_xref="taxon:77009"
/clone="bah24m12"
/tissue_type="top three leaves"
/dev_stage="adult, heading stage"
/clone_lib="K. Sato unpublished cDNA library, strain H602 adult, heading stage top three leaves"

ORIGIN

Alignment Scores:
Pred. No.: 2.34e-91 Length: 674
Score: 984.00 Matches: 193
Percent Similarity: 91.93% Conservative: 12
Best Local Similarity: 86.55% Mismatches: 17
Query Match: 35.41% Indels: 1
DB: 1 Gaps: 0

US-10-051-909-36 (1-553) x AV939950 (1-674)

QY 53 GluSerGlySerAspHisAspGlyValLeuArgArgPro-LeuLeuAsnThrGlySerTr 72
Db 3 GAGAGCGGGAGCGACCATGACACGCGGNAAGCAGCCGCTGTGTTCAACACGGGAGCTG 62
QY 72 pTyrArgMetSerSerArgGlnSerSerPheAlaProGlyThrSerSerMetAlaValle 92

Db 63 GTACCGGATGGGTCGGCGCAGTCCAGCCTCAGCGGGGACCTCTCCATGGCCATCAT 122
QY 92 uArgGluSerHisValSerAlaPheLeuCysThrLeuIleValAlaLeuGlyProIleG1 112
Db 123 GCGGGAGTCCCACGTCCTCGCCTTCTCTGCACCATGATCGTCGGCTCGGCCCATCCA 182
QY 112 nPheGlyPheThrSerGlyPheSerSerProThrGlnAspAlaMetValArgAspLeuAs 132
Db 183 GTTCGGCTTCACCGCGGCTTCTCTCCCGCCAGGACGCCATCATCCGACCTCAA 242
QY 132 nLeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyGlyMetValG1 152
Db 243 CCTCTCCATCTCCGAGTCTCCGTGTTCCGGCTCGGTGTCCAACGTGCGGCCCATGGTCGG 302
QY 152 yAlaIleAlaSerGlyGlnMetAlaGluTyrIleGlyArgLysGlySerLeuMetIleA1 172
Db 303 GGCCATGCCAGCGGCAGATGGCCGAGCAGATGGCCGCAAGGGTCACTGATGATTGC 362
QY 172 aAlaIleProAsnIleIleGlyTrpLeuAlaIleSerPheAlaLysAspAlaSerPheLe 192
Db 363 TGCTATTCTTAACATCATCGGCTGGCTGGCCATCTCTTCGCCAAAGACACTTCTTTCT 422
QY 192 uTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTyrThrValProVa 212
Db 423 GTATATGGGACGATGCTCGAAGGATTTGGTGTGGTGTGTCATATCTACACGGTGCCAGT 482
QY 212 lTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLe 232
Db 483 ATACATAGCAGAGATTCTCTCTCAGAACATGAGAGGCGCTCTAGGCTCTGTGAACCAAGT 542
QY 232 uSerValThrPheGlyIlePheLeuAlaTyrLeuLeuGlyMetPheIleProTrpArgLe 252
Db 543 GTCGGTAACGATTGGTATCGTGTGGCTTACATCTCGGCATGTTGTCTCTTGAGGAT 602
QY 252 uLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPhePheIlePr 272
Db 603 GCTTGCAGTGATAGGAATCTTGCCATGCAATATGATACCTCGTCTGTCTTCAATTCC 662
QY 272 oGluSer 274
Db 663 CGAATCT 669

RESULT 9
CA262285
LOCUS
DEFINITION SCEQLB2019E02.g LB2 Saccharum officinarum cDNA clone SCEQLB2019E02
5', mRNA sequence. 714 bp mRNA linear EST 26-SEP-2003
ACCESSION CA262285
VERSION CA262285.1 GI:35954083
KEYWORDS EST.
SOURCE Saccharum officinarum
ORGANISM Saccharum officinarum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
complex.
REFERENCE 1 (bases 1 to 714)
AUTHORS Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
TITLE The libraries that made SUCEST
JOURNAL Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
COMMENT Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bcccenter.fcav.unesp.br
Plate: 019 row: E column: 02
Seq primer: T7 Promoter Primer.
Location/Qualifiers

FEATURES

source
1. 714
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCEQLB2019E02"
/lab_host="DH10B"
/clone_lib="LB2"
/note="Organ: Lateral buds from plants adult plants
growing in greenhouse; Vector: pSport1; Site_1: SalI;
Site_2: NotI; An unidirectional cDNA library generated
from [lateral buds from plants adult plants growing in
greenhouse]. cDNA was prepared from polyA+ mRNA using
SuperScript Plasmid System Kit (Invitrogen). The
double-strand cDNAs were fractionated in a sepharose
CL-2B 40cm-columns and fragments sizing between 0.8 and
1.5 Kb were directionally cloned into the vector. Details
of each source of RNA and library construction can be
obtained at http://sucet.lad.ic.unicamp.br/public"

Alignment Scores:
Pred. No.: 1.74e-90 Length: 714
Score: 976.00 Matches: 188
Percent Similarity: 90.52% Conservative: 22
Best Local Similarity: 81.03% Mismatches: 22
Query Match: 35.12% Indels: 0
DB: 6 Gaps: 0

US-10-051-909-36 (1-553) x CA262285 (1-714)

QY 92 LeuArgGluSerHisValSerAlaPheLeuCysThrLeuIleValAlaLeuGlyProIle 111
Db 15 CTCCGCGACGTCCGCATCTCGGCCACGTTCTGACGCTCATTTGGGCTCGAGGTCCCATC 74
QY 112 GlnPheGlyPheThrSerGlyPheSerProThrGlnAspAlaMetValArgAspLeu 131
Db 75 CAGTTCGGTTTACCAGCGGGTACTCTCCCTCCCGACGAGGACGCCATCATTTGCTGACCTC 134
QY 132 AsnLeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyGlyMetVal 151
Db 135 GGCCTCTCCCTCTCCGAGTCTCTCGCTCTTCGGATCGTTATCAACGATAGGGCAATGGTA 194
QY 152 GlyAlaIleAlaSerGlyGlnMetAlaGluTyrIleGlyArgLysGlySerLeuMetIle 171
Db 195 GGTGCCATCTCCAGTGGCAACTTGCAGAGTATATCGGTGCGCAAGGGTCTCTCATGATC 254
QY 172 AlaAlaIleProAsnIleIleGlyTrpLeuAlaIleSerPheAlaLysAspAlaSerPhe 191
Db 255 GCTGCAATTCCAAACATAATTGGGTGGCTCGGATATCATTCGCAAAAGATTCTCTTTC 314
QY 192 LeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTyrThrValPro 211
Db 315 TTGTTTATGGTTCGGCTGCTAGAGGATTGGAGTTCGGTGTATATCTTATACAGTACCG 374
QY 212 ValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGln 231
Db 375 GTTTATATTGCAGAAATCGCTCTCAAGATCAGAGGGGAGTTCTTGGTCTGTCAATCAG 434
QY 232 LeuSerValThrPheGlyIlePheLeuAlaTyrLeuLeuGlyMetPheIleProTrpArg 251
Db 435 CTCTCCGTACAAATTGGTATATTGCTTGCCTACCTGTTTGGCATGTTTGGCTGGAGA 494
QY 252 LeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPhePheIle 271
Db 495 ATTCTTGTCTTAGGCGTTTACCTTGTTCATATACTGATTCCTGGACTGTTCTTTGTC 554
QY 272 ProGluSerProArgTrpLeuAlaLysMetAsnLeuThrGluAspCysGluThrSerLeu 291
Db 555 CCTGAATCCCAAGGTGGCTGGCAAAATGGGGAAGATGGAGGATTTTGAATATTCACTG 614
QY 292 GlnValLeuArgGlyPheGluThrAspIleThrThrGluValAsnAspIleLysArgAla 311
Db 615 CAAGTTCTCGAGGATTTTCAGACGACATCACACAGAGTAAGTAAATGAATAAAGAGATCA 674

QY 312 ValAlaSerSerSerLysArgThrThrThrSerPhe 323
Db 675 GTAGCATCATCAAGGAAGAGGACCAACCATATGTTT 710

RESULT 10
AV913420
LOCUS
DEFINITION AV913420 K. Sato unpublished cDNA library, cv. Haruna Niho
germination shoots Hordeum vulgare subsp. vulgare cDNA clone
bags22a03 5', mRNA sequence.

ACCESSION
AV913420
VERSION AV913420.1 GI:18209197
KEYWORDS EST.
SOURCE Hordeum vulgare subsp. vulgare
ORGANISM Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Hordeum.
1 (bases 1 to 681)
Sato, K., Saisho, D. and Takeda, K.
Barley EST sequencing project in NIG and Okayama Univ
Unpublished (2002)
Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
1. .681
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="Haruna Niho"
/sub_species="vulgare"
/db_xref="taxon:112509"
/clone="bags22a03"
/tissue_type="shoots"
/dev_stage="germination"
/clone_lib="K. Sato unpublished cDNA library, cv. Haruna
Niho germination shoots"

FEATURES
source
1. .681
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="Haruna Niho"
/sub_species="vulgare"
/db_xref="taxon:112509"
/clone="bags22a03"
/tissue_type="shoots"
/dev_stage="germination"
/clone_lib="K. Sato unpublished cDNA library, cv. Haruna
Niho germination shoots"

ORIGIN
Alignment Scores:
Pred. No.: 2.05e-90 Length: 681
Score: 975.00 Matches: 191
Percent Similarity: 91.89% Conservative: 13
Best Local Similarity: 86.04% Mismatches: 16
Query Match: 35.08% Indels: 2
DB: 1 Gaps: 2

US-10-051-909-36 (1-553) x AV913420 (1-681)

QY 53 GluSerGlySerAspHisAspGlyValLeuArgArgProLeuLeu--AsnThrGlySer 71
Db 18 GAGGGAGAGAGACCATGAC---ACGCGAAGCAGCGCTGCTGGTCAACACGGGGAGC 74

QY 72 TrpTyArgMetSerSerArgGlnSerSerPheAlaProGlyThrSerSerMetAlaVal 91
Db 75 TGGTACCGGATGGGTGCGGCCAGTCCAGCCTCACGGCGGGCACCTCTCCATGGCCATC 134

QY 92 LeuArgGluSerHisValSerAlaPheLeuCysThrLeuIleValAlaLeuGlyProIle 111
Db 135 ATGCGGGAGTCCACAGTCTCCGCCCTCTCTGCACCATGATCGTGGCTCGGCCCATC 194

QY 112 GlnPheGlyPheThrSerGlyPheSerSerProThrGlnAspAlaMetValArgAspLeu 131
Db 195 CAGTTCGGCTTACCGGGGGCTTCTCTCCCCACCCAGGACGCATCATCCGGCACCTC 254

QY 132 AsnLeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyGlyMetVal 151
Db 255 AACCTCTCCATCTCCGAGTTCTCCGTTGTCGGCTCGTGTCCAACGTCGGCGCCATGGTC 314

QY 152 GlyAlaIleAlaSerGlyGlnMetAlaGluTyrlleGlyArgLysGlySerleuMetIle 171
Db 315 GCGCCATCGCCAGCGCGCAGATGGCCGAGCACATGGCCGCAAGGGTCACTGATGATT 374

QY 172 AlaAlaIleProAsnIleleGlyTrpLeuAlaIleSerPheAlaLysAspAlaSerPhe 191
Db 375 GCTGCTATTCTAACATCATCGGCTGGCTGGCCATCTCTTCGCAAAAGACACCTCTTT 434

QY 192 LeuTyrlMetGlyArgLeuLeuGluGlyPheGlyValGlyIleleSerTyrlThrValPro 211
Db 435 CTGTATATGGGACGATTGCTTCAAGGATTGGTGTGGTGTCATATCTACACGGTGCCA 494

QY 212 ValTyrlleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGln 231
Db 495 GTATACATAGCAGAGATTCTCTCAGAACATGAGAGCGCTCTAGGCTCTGTGACCCAG 554

QY 232 LeuSerValThrPheGlyIlePheLeuAlaTyrlleLeuLeuGlyMetPheIleProTrpArg 251
Db 555 TTGTCGGTAACGATTGGTATCGTGTGGCTACATTCTCGGCATGTTTGTCTTGGAGG 614

QY 252 LeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPheIle 271
Db 615 ATGCTTGCAGTGATAGGAATCTTGCCATGCACATATTGATACCTGGTCTGTCTTCTT 674

QY 272 ProGlu 273
Db 675 CCCGAA 680

RESULT 11
CR291514
LOCUS
DEFINITION CR291514 Oryza sativa library (Han B) Oryza sativa cDNA clone
y745ellp5, mRNA sequence.

ACCESSION
CR291514
VERSION CR291514.1 GI:44678080
KEYWORDS EST.
SOURCE Oryza sativa
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1 (bases 1 to 718)
Han, B., Peng, Q., Huang, Y.C., Ying, K., Li, Y., Guan, J.P., Zhu, J.J.,
Zhao, Q., Hu, X., Liu, Y.L., Mu, J., Yu, Z., Chen, L., Fan, D.L.,
Weng, Q.J., Zhang, L., Lu, Y.Q., Yu, S.L., Liu, X.H., Lu, T.T.,
Zhang, Y.J., Lu, Y., Li, C., Li, T., Zhang, Y., Hu, H., Jia, P.X.,
Zhang, L., Lan, L.F., Chen, W., Wu, S.A. and Xue, Y.B.
Rice cDNA EST clone
Unpublished (2003)
Contact: Han Bin
National Center for Gene Research
Chinese Academy of Sciences
500# Cao Bao Road, Shanghai 200233, China
Email: bhan@ncgr.ac.cn
Clone requests: bhan@ncgr.ac.cn
This is rice cDNA est clone
Web site: http://www.ncgr.ac.cn.

FEATURES
source
1. .718
/organism="Oryza sativa"
/mol_type="mRNA"
/db_xref="taxon:4530"
/clone="y745ellp5"
/clone_lib="Oryza sativa library (Han B)"

ORIGIN
Alignment Scores:
Pred. No.: 4.43e-89 Length: 718
Score: 962.50 Matches: 198
Percent Similarity: 93.42% Conservative: 15
Best Local Similarity: 86.84% Mismatches: 15
Query Match: 34.63% Indels: 3

DB: 7 Gaps: 0

US-10-051-909-36 (1-553) x CR291514 (1-718)

QY 160 AlaGluTyrIleGlyArgGlySerLeuMetIleAlaAlaIleProAsnIleIleGly 179
 Db 8 GCGAGCTACATTGGCGCAAGGTCA-TTGATGATTGCTGCAATTCCAAACATCATTTGGT 66

QY 180 TrpLeuAlaIleSerPheAlaLysAspAlaSerPheLeuTyrMetGlyArgLeuLeuGlu 199
 Db 67 TGGCTTGCCATCTCCTTTGCAAAAGACTCATCGTTCTTTATATGGGACGATTGCTCGAG 126

QY 200 GlyPheGlyValGlyIleIleSerTyrThrValProValTyrIleAlaGluIleSerPro 219
 Db 127 GGGTTTGGTGTGGTGTGTCATCTCTTATACGGTGCCAGTTTACATAGCAGAAATATCACCT 186

QY 220 GlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIlePhe 239
 Db 187 CAAAACATGAGAGGTCCTTGGCTCAGTGAATCAGTTATCTGTAAACCGTTGGTATATTG 246

QY 240 LeuAlaTyrLeuLeuGlyMetPheIleProTyrArgLeuLeuAlaValIleGlyAlaLeu 259
 Db 247 TTGGCATATTTGCTCGGCATGTTTGTTCCTTGGAGGCTTCTTGTGTAATAGGAATCTTG 306

QY 260 ProCysThrMetLeuIleProGlyLeuPhePheIleProGluSerProArgTyrLeuAla 279
 Db 307 CCTTGCACTGTGTGATACCTGGCTATTCTTCATTCAGATCCCAAGATGGTTGGCA 366

QY 280 LysMetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPheGluThr 299
 Db 367 AAGATGAACATGATGATGATTTTGGAGACTTCTTACAAAGTTCTGAGGGGATTTGAGACT 426

QY 300 AspIleThrThrGluValAsnAspIleLysArgAlaValAlaSerSerLysArgThr 319
 Db 427 GACATCAGCGCGGAAGTGAATGATATAAAGAGAGCAGTAGCTTCAGCAAAACAAAGGACA 486

QY 320 ThrIleSerPheGlnGluLeuAsnGlnLysLysTyrArgThrProLeuLeuGlyIle 339
 Db 487 ACGATCCGTTTTCAGAAATTAAACCAGAAAGAAATACCGCACACCCCTAATACTAGGAATT 546

QY 340 GlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSer 359
 Db 547 GGCCTACTTGTACTGCAACAGGTAAGTGAATCACCGGAATATTGTTATGACAGGTAGC 606

QY 360 IlePheLysAlaAlaGlyValThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIle 379
 Db 607 ATCTTCCAAGCAGCAGGTCTCACAAACAGTACCTTGGGTACATGTGCACCTGGGGGT-ATC 665

QY 380 GlnValLeuAlaThrGlyValThr 387
 Db 666 CAGGTTCT-GCTACAGGAGTTACA 688

RESULT 12
 AJ794429
 LOCUS
 DEFINITION AJ794429 Antirrhinum majus whole plant Antirrhinum majus cDNA clone
 018_3_05_d17, mRNA sequence.

ACCESSION AJ794429
 VERSION AJ794429.1 GI:51109757
 KEYWORDS EST.
 SOURCE Antirrhinum majus (snapdragon)

ORGANISM Antirrhinum majus
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; lamiales; Plantaginaceae; Antirrhineae;
 Antirrhinum.

REFERENCE 1 (bases 1 to 762)
 AUTHORS Zachgo,S., Stueber,K., Saedler,H., Sommer,H. and Schwarz-Sommer,Z.
 TITLE Antirrhinum EST collection
 JOURNAL Unpublished (2003)
 COMMENT Contact: Schwarz-Sommer Z
 Molekulare Pflanzen-genetik
 MPI fuer Zuechtungs-forschung

FEATURES
 source
 1..762
 /organism="Antirrhinum majus"
 /mol_type="mRNA"
 /db_xref="taxon:4151"
 /clone="018_3_05_d17"
 /tissue_type="whole plant"
 /clone_lib="Antirrhinum majus whole plant"

ORIGIN

Alignment Scores:
 Pred. No.: 4.87e-89 Length: 762
 Score: 962.50 Matches: 189
 Percent Similarity: 86.00% Conservative: 26
 Best Local Similarity: 75.60% Mismatches: 34
 Query Match: 34.63% Indels: 1
 DB: 1 Gaps: 1

US-10-051-909-36 (1-553) x AJ794429 (1-762)

QY 88 SerMetAlaValLeuArgGluSerHisValSerAlaPheLeuCysThrLeuIleValAla 107
 Db 13 TCTTCTCAAGCTTTAGGACAGTTCTGTTTCTGTTGGCTTGTGTTAATTGTGGCT 72

QY 108 LeuGlyProIleGlnPheGlyPheThrSerGlyPheSerSerProThrGlnAspAlaMet 127
 Db 73 TTGGTCTCTATCCAGTTTGGTTTTCACAAAGTGGTTATTCTTCCCGACTCAAACTGCTATT 132

QY 128 ValArgAspLeuAsnLeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnVal 147
 Db 133 GTCAATGATCTTAAACTCACAGTTTCCAGTCTCGCTATTGTTGTCGAAACGTT 192

QY 148 GlyGlyMetValGlyAlaIleAlaSerGlyGlnMetAlaGluTyrIleGlyArgLysGly 167
 Db 193 GGAGCCATGGTTGGAGCAATAGCTAGTGGTCAGATTGCTGAGTACATAGGGAGAAAGGG 252

QY 168 SerLeuMetIleAlaAlaIleProAsnIleIleGlyTyrLeuAlaIleSerPheAlaLys 187
 Db 253 TCTTTAATGATCGTGCCATACCTAATATCATTTGGTTGGCTTGGCTTTCGTTTGCACAGA 312

QY 188 AspAlaSerPheLeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSer 207
 Db 313 GACATCTCATTTCTGTACATGGGAAGATTGTTGGAAGGATTGCGTCGGCATAATCTCT 372

QY 208 TyrThrValProValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGly 227
 Db 373 TATACTGTCTGTATATAGCTGAGATAGACACCTGAAATCTAAGGGGGCTCTGGGG 432

QY 228 SerValAsnGlnLeuSerValThrPheGlyIlePheLeuAlaTyrLeuLeuGlyMetPhe 247
 Db 433 TCAGTAAACCCAGCTCTCTGTGACGATCGGATCATGCTAGCATATTACTTGGACTATT 492

QY 248 IleProTyrArgLeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGly 267
 Db 493 TGCAATTGGAGATTGCTTGTCTAGGAATATTGCCCTTGTCTGATATTGATACCTGGC 552

QY 268 LeuPhePheIleProGluSerProArgTyrLeuAlaLysMetAsnLeuThrGluAspCys 287
 Db 553 CTCTTTTTCATCCAGAAATCTCCTAGATGGTTGGCCAAATGGGGATGACAGAAATTT 612

QY 288 GluThrSerLeuGlnValLeuArgGlyPheGluThrAspIleThrGluValAsnAsp 307
 Db 613 GAAGCCTCTCTCAAGTTCTTCCGGGGGTTGACCGTGTATATTGAGTATGAGTATCTGAG 672

QY 308 IleLysArgAlaValAlaSerSerSerLysArgThr---ThrIleSerPheGlnGluLeu 326
 Db 673 ATAAAGAAATCTGTGCTTCAACGACGACGACGGACGGGAGCATCCGCTTTCGCCGATCTT 732

QY 327 AsnGlnLysLysTyrArgThrProLeuLeu 336
 Db 733 AAGTTGAAAGAGATACTGGTTACCGTTGATG 762


```
RESULT 13
CD871121
LOCUS
DEFINITION
CD871121
CD871121
CD871121.1
EST.
CD871121
CD871121.1
GI:32554937
Triticum aestivum (bread wheat)
Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.
1 (bases 1 to 701)
Genoplante.
Genoplante, a major partnership french program in plant genomics
Unpublished (2003)
Contact: Genoplante
Genoplante
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french
plant genomics programme 'Genoplante' (http://www.genoplante.com
and http://genoplante-info.infobiogen.fr).
FEATURES
source
1..701
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="recital"
/db_xref="taxon:4565"
/clone="AZO2117H06"
/tissue_type="root"
/clone_lib="AZO2"
ORIGIN
Alignment Scores:
Pred. No.: 6.28e-86 Length: 701
Score: 932.00 Matches: 182
Percent Similarity: 90.48% Conservative: 27
Best Local Similarity: 78.79% Mismatches: 22
Query Match: 33.54% Indels: 0
DB: 6 Gaps: 0
US-10-051-909-36 (1-553) x CD871121 (1-701)
QY 285 GluAspCysGluThrSerLeuGlnValLeuArgGlyPheGluThrAspIleThrThrGlu 304
Db 7 GAGGACTTGAACCTCTCTACAAAGTTCTGAGGGGATTGAGACTGACATCACCCTCAGAA 66
QY 305 ValAsnAspIleLysArgAlaValAlaSerSerSerLysArgThrThrIleSerPheGln 324
Db 67 GTGAATGATATAAAGAGAGCGAGTAATATCAGCAACAAAGGGCGGATCCGTTTCCAA 126
QY 325 GluLeuAsnGlnLysLysTyArgThrProLeuLeuLeuGlyIleGlyLeuLeuValLeu 344
Db 127 GAGTTAAACCAAGAAATTCGCGATGCCCTGATTCTAGGAATTGGCCTGCTGTTCTA 186
QY 345 GlnAsnLeuSerGlyIleAsnGlyValLeuPheTyAlaSerSerIlePheLysAlaAla 364
Db 187 CAACAGCTAAGCGGAATCAACGCTATACTGTTCTATGCAAGTAGCATCTTCAAGCTGCA 246
QY 365 GlyValThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThr 384
Db 247 GGTATTACAAACAGTGACTTGGCCACATCTGGACTTGGAGGTATTACAGGTTCTTGCCACT 306
QY 385 GlyValThrThrTrpLeuLeuAspArgAlaGlyArgIleLeuLeuIleSerThr 404
Db 307 CTAGTTACAACCTGGTTACTAGACAGGGCTGGCCGGCTATCCTACTATTATCTTCT 366
QY 405 SerGlyMetThrLeuCysLeuLeuAlaValSerValPhePheLeuLysAspAsnIle 424
Db 367 GCTGGGATGACTATAAGCCCTTCTTGGGTTGCGGTTCGCGTCATATTTTATCAAGGACACTGTT 426
```

```
QY 425 SerGlnAspSerAsnSerTyTyIleLeuThrMetIleSerIleuValGlyIleValSer 444
Db 427 TCACAAGACTCTCACATGATTATACATATTGAGCATGGTCTCCTTGCTTGTCTATTGGCT 486
QY 445 PheValIleThrPheSerPheGlyMetGlyAlaIleProTrpLeuMetMetSerGluIle 464
Db 487 TATGTTATCGCCTTCTCCTTCGGTATGGGGCCATTCCATGGGTCATAATGCTTGAGATT 546
QY 465 LeuProValSerIleLysSerLeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeuThr 484
Db 547 CTCCCGGTGAGCATCAAGAGTCTCGCGGAAGCTTCGCGAGCTGCCCACTGGCTGACT 606
QY 485 SerPheAlaIleThrMetThrThrAsnLeuMetLeuThrTrpSerValGlyThrPhe 504
Db 607 TCCTTTGGGATAACGATGACAGCAAACTTGCTGCTCAGCTGGAGTGCTGGAGGTACATT 666
QY 505 LeuSerTyMetValValSerAlaPheThrIle 515
Db 667 GTGCTTACATGCTCGTGAGCGGCTTCACGCTC 699
RESULT 14
CNSOABHU
LOCUS
DEFINITION
CNSOABHU
1675 bp mRNA linear HTC 06-FEB-2004
Arabidopsis thaliana Full-length cDNA Complete sequence from clone
GSLTSL24ZD01 of Adult vegetative tissue of strain col-0 of
Arabidopsis thaliana (thale cress).
ACCESSION
BX819759
VERSION
BX819759.1 GI:42467369
KEYWORDS
HTC; GSLT cDNA.
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 1675)
REFERENCE
1 Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C.,
Menard,M., Cruaud,C., Querier,F., Scarpelli,C., Schachter,V.,
Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
Unpublished
2 (bases 1 to 1675)
REFERENCE
Genoscope.
AUTHORS
Direct Submission
JOURNAL
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque\_Projet\_EF/Full\_length
http://www.genoscope.cns.fr/cgi-bin/ggb?source=Arabidopsis.
FEATURES
Location/Qualifiers
1..1675
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Col-0"
/db_xref="taxon:3702"
/clone="GSLTSL24ZD01"
/tissue_type="Adult vegetative tissue"
/plasmid="pCMVSPORT_6"
1..1675
/gene="At2g48020"
gene
```


Pred. No.: 3.88e-85 Length: 769
Score: 925.00 Matches: 177
Percent Similarity: 85.66% Conservative: 38
Best Local Similarity: 70.52% Mismatches: 36
Query Match: 33.29% Indels: 0
DB: 4 Gaps: 0

US-10-051-909-36 (1-553) x BI933646 (1-769)

QY	181	LeuAlaIleSerPheAlaLysAspAlaSerPheLeuTyrMetGlyArgLeuLeuGly	200
Db	1	CTTTCCATCTCGTTTGCCAAAGATCTGCTTCTTATACATGGGAAGATTATTGGAAGGT	60
QY	201	PheGlyValGlyIleIleSerTyrThrValProValTyrIleAlaGluIleSerProGln	220
Db	61	TTTGGAGTCGGCATAATATCTTACACGGTTCTGTATATATTGCCGAGATAGCACCTCAG	120
QY	221	AsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIlePheLeu	240
Db	121	AACCTGAGAGGGGCCCTGGGGTCAGTTAACCACTCTCTGTTACAATTGGGATCATGTTG	180
QY	241	AlaTyrLeuLeuGlyMetPheIleProTyrArgLeuLeuAlaValIleGlyAlaLeuPro	260
Db	181	GCITATTGCTAGGACTTTTGTAAATTGGAGAGTCGCTTTTCTTGGAAACATTGCC	240
QY	261	CysThrMetLeuIleProGlyLeuPhePheIleProGluSerProArgTyrLeuAlaLys	280
Db	241	TGCCTTGCAATTGATACCTGGCCTATTTTTTCATCCCAAGATCTCCTCGGTGGCCAAAG	300
QY	281	MetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPheGluThrAsp	300
Db	301	ATGGGTCTGACAGATGATTTTGAAACCTCTTTGCAAGTTCTCCGAGGGTTCGATGCTGAC	360
QY	301	IleThrThrGluValAsnAspIleLysArgAlaValAlaSerSerLysArgThrThr	320
Db	361	ATTTCGGTCGAAGTAAATGAATAAAGAGGGCTGTAGCATCCACAAGCCGAAAGTCAACA	420
QY	321	IleSerPheGlnGluLeuAsnGlnLysLysTyrArgThrProLeuLeuGlyIleGly	340
Db	421	ATACGTTTTCAGATCTCAACAAGAAGATATTGGCTGCTCTCATGATAGGCATTGGA	480
QY	341	LeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValIleuPheTyrAlaSerSerIle	360
Db	481	CTGCTGTCTCTACAACAACCTCAGCGGAACCAATGGTGTGATCTTCTATTCCAGTAACATT	540
QY	361	PheLysAlaAlaGlyValThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIleGln	380
Db	541	TTCTATCGCCCGGGATTTCCTCAAGTGATGCTGCAACTTTAGGTTTGGTGTCTATCCAG	600
QY	381	ValLeuAlaThrGlyValThrThrTrpLeuLeuAspArgAlaGlyArgArgIleLeuLeu	400
Db	601	GTGTTGCCACTGCTGTTTCTACATGGCTGGTGGAATAAACTGGCCGTAGGCTTTACTG	660
QY	401	IleIleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSerValValPhePheLeu	420
Db	661	ATTGTCTGTCAGCTGGAATGGCTGGTAGTCTCCTTATTGTTTCCATTGCAITCTTTGGT	720
QY	421	LysAspAsnIleSerGlnAspSerAsnSerTyr	431
Db	721	AAGGATTTCGTAGATGAGGATTCTACCTTCTAT	753

Search completed: October 13, 2004, 19:46:51
Job time : 4511.99 secs

This Page Blank (uspto)

...is Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2004, 11:33:18 ; Search time 62.1256 Seconds
(without alignments)
5121.591 Million cell updates/sec

Title: US-10-051-909-36
 Perfect score: 2779
 Sequence: 1 PSSSSSFRPAGKKKKKNQG.....TTFVSLSIQROLQWLPECLS 553

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

```

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0 %
                  Maximum Match 10
                  Listing first 45

```

```
Database :      UniProt_02:*
          1: uniprot_sprot:*
          2: uniprot_trembl:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	1801.5	64.8		501	2	Q9XHW3	Q9xhw3 oryza sativ
2	1757.5	63.2		490	2	Q39416	Q39416 beta vulgar
3	1746.5	62.8		488	2	Q93YP9	Q93yp9 arabidopsis
4	1734.5	62.4		487	2	Q9FRL3	Q9frl3 arabidopsis
5	1713	61.6		515	2	Q9LN48	Q9ln48 arabidopsis
6	965.5	34.7		480	2	Q9LEG2	Q9leg2 lycopersico
7	951.5	34.2		463	2	Q8LFR8	Q8lfr8 arabidopsis
8	943.5	34.0		463	2	Q9ZU87	Q9zu87 arabidopsis
9	939	33.8		482	2	Q8VZ15	Q8vz15 arabidopsis
10	930	33.5		482	2	Q8LBI9	Q8lbi9 arabidopsis
11	920.5	33.1		496	2	Q65799	Q65799 arabidopsis
12	898.5	32.3		462	2	Q8VZT3	Q8vzt3 arabidopsis
13	890	32.0		463	2	Q9MAA4	Q9maa4 arabidopsis
14	874.5	31.5		470	2	Q94KE0	Q94ke0 arabidopsis
15	868.5	31.3		478	2	Q94CI6	Q94ci6 arabidopsis
16	863	31.1		477	2	Q9SCW9	Q9scw9 arabidopsis
17	853.5	30.7		474	2	Q94CI7	Q94ci7 arabidopsis
18	851.5	30.6		425	2	Q9MA53	Q9ma53 arabidopsis
19	843.5	30.4		458	2	Q93Z80	Q93z80 arabidopsis
20	824	29.7		483	2	Q9ZVM0	Q9zvm0 arabidopsis
21	822.5	29.6		479	2	Q8RWK4	Q8rwk4 arabidopsis
22	820.5	29.5		474	2	Q04036	Q04036 arabidopsis
23	808	29.1		454	2	Q04037	Q04037 arabidopsis
24	769	27.7		468	2	Q9LTP6	Q9ltp6 arabidopsis
25	768	27.6		804	2	Q9MAA3	Q9maa3 arabidopsis
26	748.5	26.9		461	2	Q9M100	Q9m100 arabidopsis
27	737	26.5		498	2	Q7ZWH3	Q7zwh3 brachydanio
28	729	26.2		457	2	Q9M0Z9	Q9m0z9 arabidopsis
29	698.5	25.1		497	2	Q7Q712	Q7q712 anopheles g
30	693.5	25.0		339	2	Q9SCW6	Q9scw6 arabidopsis
31	691	24.9		482	2	Q8AYP6	Q8ayp6 gallus gall

Query Match 64.8%; Score 1801.5; DB 2; Length 501;
Best Local Similarity 72.2%; Pred. No. 1.3e-111;
Matches 351; Conservative 62; Mismatches 64; Indels 9; Gaps 2;

[illegible]

Db 125 YIGRKGSLMIAAIPNIIIGWLAIISFAKDSFLMGRLLLEGFGVGVISYVVPVYIAEIAEQT 184
QY 222 MRGALGSVNQLSVTFGIFLAYLLGMFIPWRLLAVIGALPCTMLIPGLFFIPESPRWLAKM 281
Db 185 MRGALGSVNQLSVTFGIFLAYLLGMFIPWRLLAVIGALPCTMLIPGLFFIPESPRWLAKM 244
QY 282 NLTEDCETSLQVLRGFETDITTEVNDIKRAVASSSKRTTISFQELNQKKYRTPLLLGIGL 341
Db 245 GKMEDFESSLQVLRGFETDIAVEVNEIKRSVQSSRRRTTIRFADIKQKRSVPLMVGIGL 304
QY 342 LVLQNLSGINGVLFYASSIFKAAGVTNSDLATCSLGAIQVLAATGVTWLLDRAGRRILLI 401
Db 305 LVLQNLSGINGVLFYASSIFKAAGVTNSDLATCSLGAIQVLAATGVTWLLDRAGRRILLI 364
QY 402 ISTSGMTLCLLAVSVVFFLKNISQDSNSYYILTMISLVGIVSFVITFSFGMGAIPWLM 461
Db 365 ISTTGMTITLVVSVSFFVKDNITNGSHLYSVMSMLSLVGLVAFVIFSLSGLGAIPWIIM 424
QY 462 SEILPVSIKSLGSIATLANWLTSFAITMTNMLTWSVGGTFLSYMVVSAFTIVFVVLW 521
Db 425 SEILPVNIKSLAGSVATLANWLTAWLITMTASLMSWSNGGTFAIYAACAGTLVFVCLW 484
QY 522 VPETKG 527
Db 485 VPETKG 490

RESULT 2

Q39416
ID Q39416 PRELIMINARY; PRT; 490 AA.
AC Q39416;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Integral membrane protein.
OS Beta vulgaris (Sugar beet).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllales; Amaranthaceae; Beta.
OX NCBI_TaxID=161934;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96351183; PubMed=8742332;
RA Chiou T.-J., Bush D.R.;
RT "Molecular cloning, immunochemical localization to the vacuole, and
RT expression in transgenic yeast and tobacco of a putative sugar
RT transporter from sugar beet.";
RL Plant Physiol. 110:511-520(1996).
RN [2]
RP SEQUENCE FROM N.A.
RA Chiou T.-J., Bush D.R.;
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the sugar transporter family.
DR EMBL; U43629; AAB53155.1; -.
DR PIR; T14545; T14545.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005351; F:sugar porter activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0008643; P:carbohydrate transport; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub transporter.
DR InterPro; IPR003663; Sugar_transpt.
DR InterPro; IPR005829; Sug transporter.
DR Pfam; PF00083; Sugar tr; 1.
DR PRINTS; PR00171; SUGRTRNSPORT.
DR TIGRfams; TIGR00879; SP; 1.
DR PROSITE; PS50850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN_1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Sugar transport; Transmembrane; Transport.
SQ SEQUENCE 490 AA; 53020 MW; 3F8C96D26989CBCB CRC64;

Query Match 63.2%; Score 1757.5; DB 2; Length 490;
Best Local Similarity 66.2%; Pred. NO. 1.1e-108;
Matches 337; Conservative 76; Mismatches 73; Indels 23; Gaps 3;
QY 39 MGGGSRGGAGAGEESGDHGVLRRLPLNTGWSYRMSSRQSSFAPTSSMAVLRSHVS 98
Db 1 MSSDSEAGLGGG-----GDLRKPFLHTGWSYRMGRQSSL---MGSSQVIRESSIS 49
QY 99 AFLCTLIVALGPITQFTSGFSSPTQDAMVRDLNLSISEFSAFGLSNVGMVGAIASGQ 158
Db 50 VLACVLIVALGPITQFTAGYSSPTQSAITNELGLSVAEYSWFGSLNVGMVGAIASGQ 109
QY 159 MAEYIGRKGLMIAAIPNIIIGWLAIISFAKDSFLYMGRLLEGFGVGIIISYTPVYIAEIS 218
Db 110 ISEYIGRKGLMIAAIPNIIIGWLAIISFAKDSFLYMGRLLEGFGVGIIISYTPVYISEIA 169
QY 219 PQNMRGALGSVNQLSVTFGIFLAYLLGMFIPWRLLAVIGALPCTMLIPGLFFIPESPRWL 278
Db 170 PQNLRGALGSVNQLSVTIGIMLSYMLGLFVPWRILAVLGLPCTILIPGLFFIPESPRWL 229
QY 279 AKMNLTEDCETSLQVLRGFETDITTEVNDIKRAVASSSKRTTISFQELNQKKYRTPLLLG 338
Db 230 AKMGMMEEFEVSLQVLRGFETDISLEVNEIKRSVASSSKRTTIRFAELRQRRYWLPLMIG 289
QY 339 IGLLVQLNLSGINGVLFYASSIFKAAGVTNSDLATCSLGAIQVLAATGVTWLLDRAGRR 398
Db 290 NGLLILQQLSGINGVLFYSSTIFKEAGVTSSNAATFGLGAVQVATVVTWLVLDKSGRRL 349
QY 399 LLIISTSGMTLCLLAVSVVFFLKNISQDSNSYYILTMISLVGIVSFVITFSFGMGAIPW 458
Db 350 LLIVSSSGMTLSLVVAMSFFLKEMVSDSTWYSVFSILSVGVVAMVVTFSLGIGAIPW 409
QY 459 LMMSEILPVSIKSLGSIATLANWLTSFAITMTNMLTWSVGGTFLSYMVVSAFTIVFV 518
Db 410 IIMSEILPINIKGLAGSIATLANWVFAVITMTANIMLSWNSGGTFSIYMVVCFTVAFV 469
QY 519 VLWVPETKGXNSRGDTIFVSLSIQRLQW 547
Db 470 VIWVPETKGRT-----LEEIQW 486

RESULT 3

Q93YP9
ID Q93YP9 PRELIMINARY; PRT; 488 AA.
AC Q93YP9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Similar to integral membrane protein.
GN Name=Atg19450;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC euroids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Nguyen M., Karlin-Neumann G., Southwick A., Lam B., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;


```
Db      8  BEARND---LRRPFIHTGWSYRMGSRQSSM--MGSSQVIRDSSISVLACVLIVAGLPI 60
QY     112  QFGFTSGFSSPTQDMVRDLNLSISEFSAFGSLNSVGMVGAIASQMAEYIGRKGLMI 171
Db      61  QFGFTCGYSSPTQAATKDLGLTVSEYVFGSLNSVGMVGAIASQIAEYIGRKGLMI 120
QY     172  AAIPTNIIGWLAIISPAKADASYMGRLLLEGFGVGIISYTPVPIAIEISPNMRGALGSVNQ 231
Db     121  AAIPTNIIGWLAIISPAKADASYMGRLLLEGFGVGIISYTPVPIAIEISPNMRGALGSVNQ 180
QY     232  LSVTFGIFLAYLLGNFIPWRLLAVIGALPCTMLIPGLFFIPESPRWLAKMNLTEDETS 291
Db     181  LSVTIGIMLAYLLGLFVPWRILAVLIGLIPCTLLIPGLFFIPESPRWLAKMNLTEDETS 240
QY     292  QVLRFETDITTEVNDIKRAVASSSKRTTISFQELNKKYRTPTLLGIGLLVLQNLGIN 351
Db     241  QVLRFETDITTEVNDIKRAVASSSKRTTISFQELNKKYRTPTLLGIGLLVLQNLGIN 300
QY     352  GVLFYASSIFKAAGVTNSDLATCSLGAIQVLATGVTTLWLLDRAGRILLIISTSGMTLCL 411
Db     301  GVLFYSSSTIFESAGVTSSNAATFGVGAIQVATAISTWLVDKAGRLLLTISVVGMTISL 360
QY     412  LAVSVVFFLKDNIQSDNSYYILTMISLVGIVSVFVITFSFGMGALPWLMMSEILPVSIKS 471
Db     361  VIVAAAFYLKEFVSPDSMDYSLWSILSVGVVAMVVFVFLSGMGPIPLIMSEILPVNIK 420
QY     472  LGGSIATLANWLTSPAITMTNLMLTWSVGGFELSYMVVSFTIVFVVLVVPETKG 527
Db     421  LAGSIATLANWFFSWLTMTANLLAWSSGGFTLYGLVCAFTVVFVTLVVPETKG 476

RESULT 5
Q9LN48 ID Q9LN48 PRELIMINARY; PRT; 515 AA.
AC Q9LN48;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE F18014.22;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Khan S.,
RA Kim C., Altafi H., Bei B., Chin C., Chiou J., Choi E., Conn L.,
RA Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B., Lee J.,
RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M.,
RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,
RA Ecker J.R.;
RA Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
RA Khan S., Kim C., Altafi H., Bei B., Chin C., Chiou J., Choi E.,
RA Conn L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B.,
RA Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharsky N.,
RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,
RA Theologis A., Ecker J.;
RA Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Chao Q., Brooks S., Buehler E., Johnson-Hopson C., Khan S., Kim C.,
RA Shinn P., Altafi H., Bei B., Chin C., Chiou J., Choi E., Conn L.,
RA Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B., Lee J.,
RA Lenz C., Li J., Liu A., Liu J., Liu S., Mukharsky N., Nguyen M.,
RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
RA Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N., Theologis A.,
RA Ecker J.;
```

```
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the sugar transporter family.
DR EMBL; AC025808; AAF79445.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005351; F:sugar porter activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0008643; P:carbohydrate transport; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR003663; Sugar_transpt.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF00083; Sugar_tr; 1.
DR PRINTS; PR00171; SUGTRNSPORT.
DR TIGRFAMs; TIGR00879; SP; 1.
DR PROSITE; PS50850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN_1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Sugar transporter; Transmembrane; Transport.
SQ SEQUENCE 515 AA; 55930 MW; 80574480E64EF8D4 CRC64;

Query Match 61.6%; Score 1713; DB 2; Length 515;
Best Local Similarity 66.4%; Pred. No. 1e-105;
Matches 334; Conservative 67; Mismatches 68; Indels 34; Gaps 4;

QY     52  BEGSDHGDGVLRRPLNTGWSYRMSSRQSPAPGTSSMAVLRESHVSAFLCTLIIVAGLPI 111
Db      9  BEGRND---LRRPFLHTGWSYRMGSRQSSM---LESSQVIRDSSISVLACVLIVAGLPI 61
QY     112  QFGFTSGFSSPTQDMVRDLNLSISEFSAFGSLNSVGMVGAIASQMAEYIGRKGLMI 171
Db     62  QFGFTCGYSSPTQAATKDLGLTVSEYVFGSLNSVGMVGAIASQIAEYVGRKGLMI 121
QY     172  AAIPTNIIGWLAIISPAKADASYMGRLLLEGFGVGIISYTPVPIAIEISPNMRGALGSVNQ 231
Db     122  AAIPTNIIGWLAIISPAKADASYMGRLLLEGFGVGIISYTPVPIAIEISPNMRGALGSVNQ 181
QY     232  LSVTFGIFLAYLLGNFIPWRLLAVIGALPCTMLIPGLFFIPESPRWLCLYNRIAYIVEF 241
Db     182  LSVTIGIMLAYLLGLFVPWRILAVLGLVLPCTLLIPGLFFIPESPRWLCLYNRIAYIVEF 241
QY     278  ---LAKMNLTEDETSLOVLRGFETDITTEVNDIK-----RAVASSSKRTTISFQ 324
Db     242  VALLAKMGLTDDFETSLQVLRGFETDITVEVNEIKVVTCLKKCFDRSVASSSKRSARFV 301
QY     325  ELNQKKYRTPLLGLLVLQNLSGINGVLFYASSIFKAAGVTNSDLATCSLGAIQVLAT 384
Db     302  DLKRRRYFFPLMVGIGLLALQQLGGINGVLFYSSSTIFESAGVTSSNVATFGVGVQVAT 361
QY     385  GVTTWLLDRAGRRILLIISTGMTLCLLAVSVVFFLKDNIQSDNSYYILTMISLVGIVS 444
Db     362  GIATWLVDKAGRRLLLMISSIGMTISLVIVAVAFYLFKEFVSPDSNMNLSMVSVGVVA 421
QY     445  FVITFSFGMGALPWLMMSEILPVSIKSLGGSIAATLANWLTSPAITMTNLMLTWSVGGTF 504
Db     422  MVISCSLGMGPIPLIMSEILPVNIKLAGSIATLLNWFVSLVLTMTANMLLAWSSGGTF 481
QY     505  LSYMVVSFTIVFVVLVVPETKG 527
Db     482  TLXALVCGFTVVFVSLVVPETKG 504

RESULT 6
Q9LEG2 ID Q9LEG2 PRELIMINARY; PRT; 480 AA.
AC Q9LEG2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative sugar transporter.
GN Name=st3;
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
```


